GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2004, 12:22:29 ; Search time 826 Seconds
(without alignments)
10524.263 Million cell updates/sec

Title: US-10-698-235-1
Perfect score: 1656
Sequence: 1 aagtaatatcacaatagsgg.........ggtgttkggaaraaataacg 1656
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 414886 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	292.4	17.7	110000	2	AAT42063_09	Continuation (10 o
2	197.8	11.8	1572	3	AAA15298	Aaa15298 DNA encod
3	196.2	11.8	1575	2	AAZ12105	Aaz12105 Neisseria
4	196.2	11.8	12438	3	AAA81550	Aaa81550 N. mening
5	196.2	11.8	110000	3	AAA81489_6	Continuation (7 of
6	196.2	11.8	172325	3	AAF21613	Aaf21613 Neisseria
7	196.2	11.2	349980	10	AAF21612	Aaf21612 Neisseria
8	185.8	11.2	1467	2	AAZ41628	Abz41628 N. gonorr
9	185.8	11.2	1575	2	AAZ12107	Aaz12107 Neisseria
10	185.8	11.2	1575	2	AAA15323	Aaa15323 DNA encod
11	120.8	7.3	1275	2	AAZ12106	Aaz12106 Neisseria
12	96.6	5.8	110000	10	ACF65385_0	Acf65385 Photorhab
13	96.6	5.8	110000	10	ACF67367_29	Continuation (30 o
14	95.4	5.8	1284	10	ACF69918	Acf69918 Photorhab
15	95.4	5.8	1335	9	ADA31630	Ada31630 DNA encod
16	95.4	5.5	47108	8	ABK31510	Abk31510 Signal tr
17	91.4	5.5	8056	8	ABZ10246	Abz10246 Haematopo
18	89.4	5.4	1377	11	ABD11809	Abd11809 Pseudomon
19	88.2	5.3	1383	11	ABD11702	Abd11702 Pseudomon
20	88.2	5.3	1410	11	ABD12145	Abd12145 Pseudomon
21	87.2	5.3	19380	6	AAS61427	Aas61427 Human gen

22	86.6	5.2	8079	6	ABL92313	Abl92313 Chemicall
23	86.4	5.2	7814	4	AAS46530	Aas46530 Tumour su
24	85.6	5.1	3683	8	ABZ10199	Abz10199 Haematopo
25	84.4	5.1	14147	10	ADB54225	Adb54225 Pretreate
26	84.4	5.1	14147	10	ADE84163	Ade84163 Human lym
27	84.2	5.1	6255	6	ABL32960	Abl32960 Human imm
28	83.8	5.0	8056	8	ABZ10100	Abz10100 Haematopo
29	83.6	5.0	8776	6	ABK40068	Abk40068 Human che
30	83.4	5.0	5822	6	ABL33096	Abl33096 Human imm
31	83.2	5.0	7892	6	ABK40056	Abk40056 Human che
32	82.8	5.0	6419	6	ABL32267	Abl32267 Human imm
33	82.8	5.0	113515	6	ABL34174	Abl34174 Human imm
34	82.2	5.0	8056	8	ABZ10246	Abz10246 Haematopo
35	81.8	4.9	6668	8	ABL33697	Abl33697 Human imm
36	81.8	4.9	8222	8	ACF62816	Acf62816 Colon can
37	81.8	4.9	11222	10	ADB54318	Adb54318 Pretreate
38	81.4	4.9	17848	4	AAS45323	Aas45323 Chemicall
39	81.4	4.9	17848	6	ABK39976	Abk39976 Human che
40	81.4	4.9	17848	6	ABK28164	Abk28164 DNA trans
41	81.2	4.9	3683	8	ABZ10053	Abz10053 Haematopo
42	81.2	4.9	14147	4	AAS46743	Aas46743 Tumour su
43	81.2	4.9	14147	6	ABK33955	Abk33955 Human DNA
44	81.2	4.9	14147	8	ADA20386	Ada20386 Prostate
45	81.2	4.9	14147	8	ADA84193	Ada84193 Human ren

ALIGNMENTS

RESULT 1
AAT42063_09
Continuation (10 of 19) of AAT42063 from base 900001 (Haemophilus influenzae complete ger
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

WP Fragment Name	Begin	End
AAT42063_00	1	110000
AAT42063_01	100001	210000
AAT42063_02	200001	310000
AAT42063_03	300001	410000
AAT42063_04	400001	510000
AAT42063_05	500001	610000
AAT42063_06	600001	710000
AAT42063_07	700001	810000
AAT42063_08	800001	910000
AAT42063_09	900001	1010000
AAT42063_10	1000001	1110000
AAT42063_11	1100001	1210000
AAT42063_12	1200001	1310000
AAT42063_13	1300001	1410000
AAT42063_14	1400001	1510000
AAT42063_15	1500001	1610000
AAT42063_16	1600001	1710000
AAT42063_17	1700001	1810000
AAT42063_18	1800001	1830001

Query Match   17.7%; Score 292.4; DB 2; Length 110000;
Best Local Similarity 85.1%; Pred. No. 3.3e-43;
Matches 336; Conservative 1; Mismatches 57; Indels 1; Gaps 1;

Qy   8 ATCACAATAGSGGATCCACGAGCTTCTATTAGGTATCGTATTGGCTGCAGAGGGATATCC 67
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41919 ATATCACAATGGGATCTGCGAGCTTCT-TTAGGTATCGTATTGGCTGCAGAGGGATACCC 41977

Qy   68 AAAGGATTATCGCAAAGGCGATGAAATCAGCGGATTGCCTAAAAGTGCGGTCAAAAACGA 127
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 41978 AAAAGATTATCGCAAAGGCGATGAAATCAGCGGATTGCCTAAAAGTGCGGTCAAAAACGA 42037

Qy   128 GAAAGTTTTCTTAGCGGGTGTCGCAGAACAAGAGGCAAGCTAGTCACAACGGCGGTCG 187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 42038 GAAAGTTTTCTTAGCGGGTGTCGCAGAACAAGAGGCAAGCTAGTCACAACGGCGGTCG 42097

Qy   188 TGTACTTTGTGTGACTGCGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA 247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 42098 TGTACTTTGTGTGACTGCGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA 42157

```
Qy    248 ATTGGCTGAGCAAATTCAATGGTCTGGGCGTTTTTATCGTCGAGACATTGGTTACAGGGC 307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  42158 ATTGGCTGAGCAAATTCAATGGTCTGGGCGTTTTTATCGTCGAGACATTGGTTACAGGGC 42217

Qy    308 TGTGGAACGAGAACAAGCAAATAGTAGAAATCTTGTTGAATTAATTAGATAAAAAT 367
          |||||||||||  |||||||
Db  42218 TGTGGAACGAGAGTTGCAAATAATAATCCATATCATTATGATGAATTAATTTCATTTAAGT 42277

Qy    368 ATTGTACAGGGTAGAAATTGTATTTTCCTAGGATTT 402
          |||||                    |||  ||
Db  42278 GATGAAAATTTAGGAATGAAAATTTTAATTAAATTT 42312
```

RESULT 2

```
AAA15298
ID   AAA15298  standard; DNA; 1572 BP.
XX
AC   AAA15298;
XX
DT   04-SEP-2000  (first entry)
XX
DE   DNA encoding a polypeptide of a Neisseria pathogenic strain.
XX
KW   Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX
OS   Neisseria meningitidis.
XX
FH   Key            Location/Qualifiers
FT   CDS            1..1572
FT                  /*tag= a
XX
PN   WO200026375-A2.
XX
PD   11-MAY-2000.
XX
PF   28-OCT-1999;  99WO-FR002643.
PR   30-OCT-1998;  98FR-00013693.
XX
PA   (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
PA   (INMR ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI   Aujame L, Bouchardon A, Renauld-Mongenie G, Rokbi B, Nassif X;
PI   Tinsley C, Perrin A;
XX
DR   WPI; 2000-365622/31.
DR   P-PSDB; AAY93267.
XX
PT   New polypeptide specific for pathogenic Neisseria useful in therapeutic
PT   or preventative vaccines and for diagnosis.
XX
PS   Claim 2; Page 56-59; 187pp; French.
XX
CC   The present sequence encodes a protein that is specific for pathogenic
CC   strains of Neisseria. The polynucleotides, polypeptides, or their
CC   antigenic fragments, are used in vaccines to treat or protect against
CC   Neisseria infections, particularly by N. meningitidis. The polynucleotide
CC   sequence is also used for recombinant production of the polypeptide and
CC   to produce attenuated Neisseria strains that overexpress it, or express
CC   it in a non-toxic mutant form
XX
SQ   Sequence 1572 BP; 380 A; 383 C; 395 G; 414 T; 0 U; 0 Other;
```

```
Query Match             11.9%;  Score 197.8;  DB 3;  Length 1572;
Best Local Similarity   51.6%;
Matches  533;  Conservative    2;  Mismatches  480;  Indels   18;  Gaps    3;
```

```
Qy    589 CTGAAAGTTATTGTCTCGTTGTTGTTTGTTTGTTTTTTTCAAAATATAGAA 648
          | ||||| ||  |||            |||||||  | |||||  |
Db    110 CAGAAACGTTGTCGCTGACATTAGATTGCTGCCTGTATCTGTTTGCGCGTTATAAGG 169

Qy    649 TTACAAGAATTATTAGTTGGCGCTTTATTTGCAATAAGTATTGTTGTTAATAAGTACATT 708
```

```
Db    170 CTTCGCGGCTGCTCGATTGCGTGTTTTTTCGGCGTTCAGCATTATTGCCAACACATGTACATT 229
          ||  |       ||   |||  |||||||   |||||||   |||||||||||
Qy    709 ATGCAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTTGCATTTAAAGAAATTA 768
          |||| ||| |||||| |||||| | |||||  |||| || | |||  || | |
Db    230 ATGCGGTTTATCAAAGTTGGATGACGGGCATCAATTATTGGCTGCTGATGCTGAAAGAGATTA 289

Qy    769 ATGAGATAACAACAAATGCTGGCTTAACAATGATAGATAAATTCATTATATCCATTGTTATTTG 828
Db    290 CCGAAGTCGGACAGTGCGGCGGCCGTCGATGTTGGGCGTGCTGCCTCGCGTTGTGGG 349

Qy    829 GTTTATTTGAAGTGCTGTGTTTTAAGTTTCATCAAAGGAAAAGTATATAAAC 888
Db    350 GCGTGTTGGAAGTCATGTGTTGTTTGCGACGTTCGCCAAGTTCCACGTGATGACGCATT--- 406

Qy    889 TTTCTTGGATTTTGACTTATTTTATTTTGTGATGAGTGGTTTTTGTTGAGCGT 948
Db    407 -------TTTCTGCCGATATACTGTTTGCCTTCCACTTGTGATGATGATGATTTTCGTGCGTT 457

Qy    949 ATACAACAAAATCCATGAGCGTTTTATTTCACCTAACACTGTTTATTCTCGATTAAAAT 1008
Db    458 CGTTCGACACGGAAACAAGAGCACGGTATTTCGCCCAAACCACATACAGCCGCATCAAAG 517

Qy   1009 CCAATTATTTATCGTTGGGTTATTTTATAGGACGAATTGTTCCTTATGAGATATTTCTT 1068
Db    518 CCAATTATTTCAGCTTCGGTTATTTTTGTCGGACGCGGTGTTGCCGTATCAGTTGTTTGATT 577

Qy   1069 TATCTAATATCCTCTTTATCATAAATCTAAGCCTATGAAATCGGGCTCTCCGVAAAATT 1128
Db    578 TAAGCAGGATTCCCGCTTTAAGCCTTTAAGCCTGCTCCAAGCAAAATCGGG---CAGGGCAGTG 634

Qy   1129 MARGAATATAATTTAATTAATGGGGAAGTGCGACCTCAAGTCATTTTAGTGCTTTTG 1188
Db    635 TTCAAAATATCGTCCTGATTATGGGCGAAAGCGAAAGCGCGGCGCATTTGAAGCTGTTTG 694

Qy   1189 GTTACGGGAGAAAAAACATCTCCTTTTTTAGATAG---CTTAAAATAAAATCAGGAGCTC 1245
Db    695 GCTACGGACCGGAAACTTCGCCGTTTTTAACCCGGCTGTGCAAGCCGATTTTAAGCCGA 754

Qy   1246 TTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTTCTTTACCAATGTTTTTTA 1305
Db    755 TTGTGAAACAAAGTTATTCCGGAGCTTAGCCAGTGTCCCTGCCCAGTTTTTTCA 814

Qy   1306 ATGCAATYCCTTAYCCAAATGGAATACAACAGATAGTCTAAAGGAGATACGAATTATTA 1365
Db    815 ATGCGGATACCGCCGCAACGGCTTGGAACAAATCAGCGGCGGCGATACTAATATGTYCC 874

Qy   1366 ATTTAGCGAAAGAGAGCAAGGCTTCAGACATACATATTTTTTATTCAGCTCAAGCTAGGGATGATA 1425
Db    875 GCCTCGCCAAAGAGCAGCAGGGCTATGAAACGTATTTTTACACGCGACAGGGCGAAAACGAGA 934

Qy   1426 TGCATATGATCAATTTTTAGGCAGCTTCGATTCGATTCAGCCGACCAGC 1485
Db    935 TGGCGATTTTGAACTTAATCGGTAAGAAATGGATAGACCATCTGATTCAGCCGACCAGC 994

Qy   1486 AAGGGTATTCTTTAAGAGATTCAATGCCTGATAATAAATTACTTCCTGCTTTAAAAATA 1545
Db    995 TTGGCTACGGCAACGGCGACAATATGCCCGATGAGAGCTCTCGCCGTGTTGCACAAA 1054

Qy   1546 TTAATTTAGATAATGGTTATCCATTTTGTTGTTTTTACATCATAGAGGAGTCATATTCCCT 1605
Db   1055 TCAATTTGCAGCAGGCAGGCATTTTATCCTGTTGCACCAAACGTGTTCGCACGCCCCAT 1114

Qy   1606 ATGGGGCCATTATT 1618
Db   1115 ACAGGCGCATTGTT 1127
```

RESULT 3

```
AAZ12105
ID   AAZ12105  standard; DNA; 1575 BP.
XX
AC   AAZ12105;
```

```
Db  406  -------TTTTCTGCGATATACTGTTTGCCTTCCTATGCTGTGATGATTTCGTGCGTT  457
Qy  949  ATACAACAAATCCATGACGCGTTTATTCACCTAACACTGTTTATTCTCGATAAAAT 1008
         |||||  ||||| ||||| |||  | |||| | ||||||  |||||| |||||||||
Db  458  CGTTCGACACGAAACAAGAGCACGTATTTCGCCCAAACGACATACAGCCGCATCAAAG  517

Qy 1009  CCAATTATTATCGTGGGTTATTTTATAGGACGAATTGTTCCTTATGAGATAATTTTCTT 1068
         || ||||||||| ||| || ||||||||||||||||||||||||||||||||||||||
Db  518  CCAATTATTTCAGCTTCGGTTATTTTGTCGGACGGCTGTTGCCGTATCAGTTGTTTGATT  577

Qy 1069  TATCTAATATTCCTTCTTTATCATAAATCAAGCCTATGAAATCGGGCTCTCCGTAAAATT 1128
         |||| ||||:   ||||||||||  :|  ||  | ||||||||  |   ||||:||||||:
Db  578  TAAGCAGGATTCCCGCCTTAAGCAGCTGCTCCAAGCAAAATCGGG---CAGGGCAGTG  634

Qy 1129  MARGAATATAATTTTAATTAATGGGGAAAGTGCGACCTCAAGTCATTTAGTGCTTTTG 1188
         |    |||| |||||||| ||||  ||||| | ||  ||||||||||| ||| || |||
Db  635  TTCAAAATATCGTCTCCTGATTATGGCCATTGAAAGCGGCGGCCATTTGAAGCTGTTTG  694

Qy 1189  GTTACGGGAGAAAAACATCTCCTTTTTTAGATAG---CTTAAAATATAAATCAGGAGCTC 1245
         | |||||||  |||| ||| || ||||| |||||   |||  ||| | || |||||| |
Db  695  GCTACGGACGCGAAACTTCGCCGTTTTTAACCCGCTGTCGCAAGCGATTTTAAGCCGA  754

Qy 1246  TTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACACAGATAGTTCTTCTTTACCAATGTTTTTA 1305
         ||||||||||||  ||| | ||  |||||||||||   ||:||  ||| ||||||||| ::||
Db  755  TTGTGAAACAAAGTTATTCCGCAGGCTTTATGATGCTGTCTGCTGGTCCCTGCCCAGTTTTTCA  814

Qy 1306  ATGCAATYCCTTAYCCAAATGAATGACACACAGATAGTCTAAAGACAGCAATTATTATTA 1365
         ||| |||  ||||  || || ||||| |||  | |||| |||||||||||  || ||||||
Db  815  ATGCGGATACCGCACGCCACGGCTTGGAACAAATCAGCGGCGGCGATACCAACAATATGTTCC  874

Qy 1366  ATTTAGCGAAAGAGCAAGGCTTCAGACACATATTTTTATTCAGCTCAAGCTAGGGATGATA 1425
         || || ||||||||||| | ||   | | || || | | |||    ||||| || || ||
Db  875  GCCTCCGCCAAAGAGGAGGGGCTATGAAACGTATTTTTACAGCGCGGCGGCGGAAAACGAGA  934

Qy 1426  TGCATATGATCAATTTTTAGGAGGAGCTTGGATTGATGAGATTATTCGTTTTCCAGATAATG 1485
         ||| | || ||||| ||||  |||||  ||| ||||| | |||||||||| |||||||| ||
Db  935  TGGCGATTTTGAACTTAATCGGTAAGAAATGGATGAGCCATCTGATTCAGCCGACGCAAC  994

Qy 1486  AAGGGTATTCTTTAAGAGATTCAATGCCTGATATAAAATTACTTCCGCTTTAAAAATA 1545
         || |||||| ||||| || ||| ||||| || | | || ||| ||| |:||  ||| |
Db  995  TTGGCTACGGCAACGCGACAAATGCCCGATGAGAAGTCGTGCTGCCGTGTTGTCGACAAAA 1054

Qy 1546  TTAATTAGATAATGGTTATCATTTTGTTTGTTTACATCATAGAGGGAGTCATATTCCCT 1605
         |||  || || |||| |  ||  ||||| |  |:|| ||||| |  || | |||||  ||
Db 1055  TCAATTGCAGCAGGCAAGCAAGTATCGTGTTGTCACCAACGCGTTCGCACGCGCCCAT 1114

Qy 1606  ATGGGGCATTATT 1618
         ||  ||||||||
Db 1115  ACGCGCCATTGTT 1127
```

RESULT 4
AAA81550
ID   AAA81550 standard; DNA; 12438 BP.
XX
AC   AAA81550;
XX
DT   04-DEC-2000 (first entry)
XX
DE   N. meningitidis partial DNA sequence gnm_97 SEQ ID NO:97.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW   antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW   Meningococcus B; MenB; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO20002430-A2.
XX
PD   20-APR-2000.
XX
PF   08-OCT-1999;   99WO-US023573.
XX
DT   08-OCT-1999 (first entry)
XX
DE   Neisseria meningitidis complete ORF81 sequence.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9924578-A2.
XX
PD   20-MAY-1999.
XX
PF   09-OCT-1998;   98WO-IB001665.
XX
PR   06-NOV-1997;   97GB-00023516.
PR   14-NOV-1997;   97GB-00024190.
PR   18-NOV-1997;   97GB-00024386.
PR   27-NOV-1997;   97GB-00025158.
PR   10-DEC-1997;   97GB-00026147.
PR   14-JAN-1998;   98GB-00000759.
PR   01-SEP-1998;   98GB-00019016.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
XX
DR   WPI; 1999-327407/27.
DR   P-PSDB; AAY38651.
XX
PT   Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT   diagnosis, treatment and prevention of infection.
XX
PS   Claim 9; Page 205; Page 524pp; English.
XX
CC   Nucleotide sequences AAZ11972-Z12358 represent open reading frames (ORFs)
CC   of Neisseria meningitidis and N. gonorrhoeae which encode antigenic
CC   proteins (see AAY38499-Y38944). The antigenic proteins, their fragments,
CC   their nucleic acids and antibodies are used for diagnosis, prevention (as
CC   vaccines) or treatment of Neisseria infections, such as meningitis,
CC   septicaemia and gonorrhea. Both organisms are closely related. Fragments
CC   of the nucleic acids are useful as hybridisation probes and antisense
CC   reagents
XX
SQ   Sequence 1575 BP; 376 A; 378 C; 405 G; 416 T; 0 U; 0 Other;

Query Match   11.8%; Score 196.2; DB 2; Length 1575;
Best Local Similarity 51.5%; Pred. No. 4.6e-26;
Matches   532; Conservative   2; Mismatches 481; Indels 18; Gaps   3;

```
Qy  589  CTGAAAGTATTGTCTGCGTTGTTGTTGTTGTTTGTTTGTTTTTCAAAATATAGAA  648
         ||||||||||||||||| ||||||||||||||||||||| |||| ||||
Db  110  CGGAAACGTTGCCGCTGACATTTGTGCTGCCTGCGCTGATTCGTGCTGCGGCGTTTATAAGG  169

Qy  649  TTACAAGATATATTAGTTGGCGCTTATTGCAATAAGATTAGTTGTTAATAATGTACATT  708
         ||| ||| | | | | || ||  ||| |||| | | |||  ||| | | ||| ||| ||
Db  170  TGACGCGTTTGTTGATTGCGTGGTGTTTTTGCGTGTTGCCACAATGTCATT  229

Qy  709  ATGCAGTATACCAATCTTCGGATTCGGATTCTGACCTGGTCCATTCACTTGCATTTAAAGAAAATTA  768
         ||| || | |||| |  | | || ||| ||||| | ||| | ||  | |||| |||||||||
Db  230  ACGCGGTTTATCAAAGCTGGATGACGGGCATCATTATTGGCTGACTGCTGAAAGAGGGTTA  289

Qy  769  ATGAGATAACAAAATGCTGGCTTAACAATGATAGATAAATTCAATATATCCATTGTTATTG  828
         ||| |||  | || |   ||||||   |||   | | | ||||| ||| ||| ||||||||
Db  290  CCGAAGTCGGCAGCGCGGGTGCCGTGCGTGGGTGGATAAGTTGTGGGCCTGCTGTTGTTGGG  349

Qy  829  GTTTATTTGAAGTCGTGTTTGTTTTTAAGTTTAAGTTCATAAAAAGAGAAAAGTATATAAAC  888
         |||  ||||||| || |  ||||| |||  | |||||| |||| | | |||| ||||| |
Db  350  GCGTGTTGGAAGTCATGTTTTTTGCAGCGTGATGATGTTTTTTGTTCGAGCGT  405

Qy  889  TTTCTTGGATTTTGACTTTATTTTATGCTGTGATGAGTGATGTATGTTTTTGTTCGAGCGT  948
         |||  || ||| ||| |||| | |||| | ||||| |||||||  ||||||  ||||||||
```

PR   09-OCT-1998;   98US-0103794P.
PR   30-APR-1999;   99US-0132068P.
XX
PA   (CHIR) CHIRON CORP.
XX
PI   Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI   Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI   Rappuoli R, Pizza M;
XX
DR   WPI; 2000-318079/27.
XX
PT   Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT   in the diagnosis and treatment of N. meningitidis infection and other
PT   Neisserial infections, for example, N.gonorrhoea.
XX
PS   Claim 7; Page 1489-1492; 1760pp; English.
XX
CC   The present invention describes methods of obtaining immunogenic proteins
CC   from Neisseria genomic sequences. AAA81453 to AAA82414 represent
CC   specifically claimed Neisseria meningitidis genomic DNA sequences;
CC   AAA81260 to AAB25620 and AAB25663 represent Neisseria genomic DNA
CC   sequences and their corresponding proteins; AAA81259 and
CC   AAA81304 to AAA81321 represent PCR primers used in the isolation of
CC   Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
CC   Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC   used in the exemplification of the present invention. The nucleic acid
CC   sequences, protein sequences, and antibodies against them, can be used in
CC   the manufacture of a composition. The composition can be used as a
CC   medicament (or in the manufacture of a medicament) for treating,
CC   preventing or diagnosing infection due to Neisserial bacteria. For
CC   example, some of the identified proteins could be components of vaccines
CC   against Meningococcus B; against all serotypes; and/or against all
CC   pathogenic Neisseriae. Identification of sequences from the bacterium
CC   will also facilitate production of biological probes, particularly
CC   organism-specific probes. Attempts to make efficacious Meningococcus B
CC   vaccines have failed mainly due to antigen tolerance. Multivalent
CC   vaccines have also been tried but none have successfully overcome
CC   antigenic variability. The provision of further, complete sequences may
CC   provide an opportunity to identify secreted or surface exposed proteins
CC   that may be presumed targets for the immune system and which are not
CC   antigenically variable or at least more conserved than other more
CC   variable regions
XX
SQ   Sequence 12438 BP; 3661 A; 3104 C; 2565 G; 3108 T; 0 U; 0 Other;

Query Match   11.8%; Score 196.2; DB 3; Length 12438;
Best Local Similarity 51.5%; Pred. No. 5.4e-26;
Matches 532; Conservative   2; Mismatches 481; Indels   18; Gaps   3;

Qy   589 CTGAAGTTATTGTCGTTTGTTTGTTTGTTTGTTTGTTTTCAAAATATAGAA 648
          |||||||||||||| ||||| ||||| ||| ||| | ||||||||||
Db   3006 CGGAAACGTTTGCGCTGACATTTGTGATTGCTGCGCGTATCGTTTATAAGG 3065

Qy   649 TTACAAGAATATTAGTTGCGGCCTTTATTGTTGTCAATAAGTAATGTGCATT 708
          |||| | | || || |||| ||||||||||||||| |||| |||| ||
Db   3066 TGACGGTGTTGTTGATTGCGGTGTTTTTGCGGTGTTTTTGCGGTTAATAATGTGCATT 3125

Qy   709 ATGCAGTATACCAAATCTGGACCTGTTAATTACTCACTTGCATTTAAAGAAATTA 768
          |||| | | | | |||||| |||||||||||| || || | ||||||| || |
Db   3126 ACGCGGTTATCAAAGCTGGATGACGGGGCATCAATTATTGGCTGATGGCTGAAAGAGGTTA 3185

Qy   769 ATGAGATAACAAATGCTGGC 828
          ||| | | |||||||||||
Db   3186 CCGAAGTCGGCAGCGGCGTGCGTGCGATGTGCGTCGGGTTCGGG 3245

Qy   829 GTTTATTGAAGTGCGTGTTTTATTTTTTTAAGTTAAGTTTCATAAAAGAAAGTATATAAAC 888
          |||||||||||||||||| || |||| ||||||| | | |||||||||| |||||
Db   3246 GCGTGTTGGAAGTCATGTTGTTTGCAGCCTGCCAAGTTCCGCCGTAAGACGCAT----- 3301

Qy   889 TTTCTTGGGATTTTGACTTATTTTTTATGCTGTGATGATGATGTATGTTTTTGTTCGAGCGT 948
          |||||| || |||||| | | |||||| | || | |||| ||||| | | || |||||
Db   3302 -------TTTTCTGCCGATATACTGTTTGCCTTCCTAAATGCCTGATGAATTTTCGTTGCGTT 3353

RESULT 5

AAA81489_6
Continuation (7 of 9) of AAA81489 from base 600001 (N. meningitidis partial DNA sequence
Sequence split into 9 fragments   LOCUS AAA81489   Accession Aaa81489
WP   Fragment Name   Begin   End
WP   AAA81489_0   1   110000
WP   AAA81489_1   100001   210000
WP   AAA81489_2   200001   310000
WP   AAA81489_3   300001   410000
WP   AAA81489_4   400001   510000
WP   AAA81489_5   500001   610000
WP   AAA81489_6   600001   710000
WP   AAA81489_7   700001   810000
WP   AAA81489_8   800001   837096

Query Match   11.8%; Score 196.2; DB 3; Length 110000;
Best Local Similarity 51.5%; Pred. No. 6.5e-26;
Matches 532; Conservative   2; Mismatches 481; Indels   18; Gaps   3;

Qy   589 CTGAAGTTATTGTCGTTTGTTTGTTTGTTTGTTTGTTTTCAAAATATAGAA 648
          |||||||||||||| ||||| ||||| ||| ||| | ||||||||||
Db 87957 CGGAAACGTTTGCGCTGACATTTGTGATTGCTGCGCGTATCGTTTATAAGG 88016

Qy   649 TTACAAGAATATTAGTTGCGGCCTTTATTGTTGTCAATAAGTAATGTGTACATT 708
          |||| | | || || |||| ||||||||||||||| |||| |||| |||
Db 87957

Qy   949 ATACAACAAAATCCATGAGCGTTTGTTTATTCACCTAACACTGTTTATTCTGATTAAAAT 1008
          |||| |||||
Db   3354 CGTTCGACACGAAACAAGAGCACGGTATTTGCCCAAACCGACATACAGCCGCATCAAAG 3413

Qy   1009 CCAATTATTTATCGTTGGTTATTTATAGGACGAATTGTTCCTTATGAGATATTTCTT 1068
          ||||||||||| ||||||||||| ||| |||||| | | | |||||| |||||
Db   3414 CCAATTATTTATTCAGCTTCGGTTATTTTGTCGGACGCGTGTTGTTTGTTTGATT 3473

Qy   1069 TATCTATATTCCTCTTATCATAAATCTAAGCCTGCTCCAAGCAAAATCGGG 1128
          || | || ||| ||||| |||| ||||||||||||||||||||||
Db   3474 TAAGCAGGATTCCCGCCTTTAAGCAGCCTGCTCCAAGCAAAATCGGG---CAGGGCAGTG 3530

Qy   1129 MARGAATATAATTTTAATTAATGGGGAAAGTGCGACCTCAAGTCATTTTAGTGCTTTTG 1188
          ||||| || || ||| ||| |||| ||||| | |||| ||| |||| ||||
Db   3531 TTCAAAATATGTCCTGATTATGGGCGCGAAAGCGAAAGCGCGGGCCATTTGAGCGTGTTTG 3590

Qy   1189 GTTACGGGAGAAAAACATCTCCTTTTTTAGATAG---CTTAAAATATAAATCAGGAGCTC 1245
          ||||||| ||||||| | | |||||| ||| ||| |||| | | |||| ||| ||
Db   3591 GCTACGGACGCGAAACTTCGCCGTTTTTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGA 3650

Qy   1246 TTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTTCTTTACCAATGTTTTTTA 1305
          ||||||| |||| ||||| || || |||| || || | |||||||||| ||||||
Db   3651 TTGTGAAACAAGTTATTTCCCAGGCTTATGACTGCAGTGTCCCGCCCAGTTTTTTCA 3710

Qy   1306 ATGCAATYCCTTAYCCAAATGGAATACAACAGATAGCTAAAGGAGGATACGAATTTATTTA 1365
          ||| || ||| | | ||| | |||| | || || ||| ||| || | ||||| ||
Db   3711 ATGCGGATACCGCACCGCCAACGCGTTGGAACAAATCAGCGCGGCGGCGATACCAATATGTTCC 3770

Qy   1366 ATTTAGCGAAAGAGCAAGGCTTTCAGACATATTTTATTCAGCTCAAGCTAGGGATGATA 1425
          || || || ||||||| |||| |||| ||||| ||||| || | |||| |||||
Db   3771 GCCTCGCCAAAGAGCAGGCGTATGAAACGTATTTTTACAGCGCCGCGGCGGAAAACGAGA 3830

Qy   1426 TGCATATGATGCAATTTTTAGGAGGAGCTTGGATATTTGTTTCAGATAATG 1485
          |||| |||||| || | | || ||| ||||| || | || ||| ||
Db   3831 TGGCGGATTTGAACTTAATCGGTAAGAAATGATAGACATCTGATTCAGCCGACGCAAC 3890

Qy   1486 AAGGGTATTCTTTAAGAGATTCAATGCCTGATAAATAAATTACTTCCTGCTTTTAAAAATA 1545
          |||| | || ||| ||| || | ||| |||| | |||| || || ||||| | ||||
Db   3891 TTGGCTACGGCACGCGGACATATGCCGATGAGAGCTGCTGCCGTTGTTGTCGACAAAA 3950

Qy   1546 TTAATTAGATAATGGTTATCATTTGTGTTTTACATCATAGAGGGAGTCATATTCCCT 1605
          |||| ||| | |||| || |||| |||||| |||| || | ||| | |||| |||
Db   3951 TCAATTTGCAGCACGGCAAGCCATTTTATCGTGTTGTTCACCAACGCGGTTCGCACGCCCAT 4010

Qy   1606 ATGGGGCCATTATT 1618
          |||| |||||||
Db   4011 ACGGCGCATTGTT 4023

```
XX
DT   13-MAR-2001 (first entry)
XX
DE   Neisseria meningitidis B nucleotide sequence SEQ ID NO:114.
XX
KW   Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW   diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW   ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO200066791-A1.
XX
PD   09-NOV-2000.
XX
PF   08-MAR-2000; 2000WO-US005928.
XX
PR   30-APR-1999;    99US-0132068P.
PR   08-OCT-1999;    99WO-US023573.
PR   28-FEB-2000;    2000GB-00004695.
XX
PA   (CHIR ) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC;
PI   Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI   Rappuoli R, Frazer CM, Grandi G;
XX
DR   WPI; 2000-647603/62.
XX
PT   Neisseria meningitidis B full length genome sequence and open reading
PT   frames are used to detect, treat and prevent Neisserial infections.
XX
PS   Claim 7; Appendix A; 692pp; English.
XX
CC   The present invention describes the full length genome of Neisseria
CC   meningitidis B (NMB). The sequences in AAF21607 to AAF21613
CC   represent fragments of the NMB genomic sequence, as the sequence was too
CC   long to go in a record on its own it was split into 8 sequences which
CC   overlap each other at the beginning and end of each sequence by 49980 bp
CC   (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
CC   AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
CC   AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins
CC   given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
CC   primers which are used in the exemplification of the present invention.
CC   The NMB genome and fragments from it have antibacterial activity, and can
CC   be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
CC   and/or antibodies which binds to the proteins can be used in compositions
CC   for treating or preventing infection due to Neisserial bacteria or as a
CC   diagnostic reagent for detecting the presence of Neisserial bacteria or
CC   of antibodies raised to Neisserial bacteria. Computers, computer memory,
CC   computer storage medium or computer databases can be used in a search to
CC   identify open reading frames (ORFs) or coding sequences within the NMB
CC   genome. The DNA sequences provide further opportunities to find antigenic
CC   or immunogenic proteins which are more effective in vaccines than the
CC   outer membrane proteins currently used
XX
SQ   Sequence 172325 BP; 43072 A; 47583 C; 41465 G; 40205 T; 0 U; 0 Other;
```

```
Query Match             11.8%;  Score 196.2;  DB 3;  Length 172325;
Best Local Similarity   51.5%;          Pred. No. 6.7e-26;
Matches  532;  Conservative    2;  Mismatches  481;  Indels  18;  Gaps  3;
```

```
Qy         589 CTGAAAGTTATTGTCTGTTTGTTTGTTTGTTTGTTTGTTTTTCAAAATATAGAA 648
               ||||| |||||||| ||||||||| ||||| ||| ||| |||||||||||||||
Db       23186 CGGAAACGTTTGCGCTGACATTTGTGATTGCTGCGTGTATCGTGTTTGCGGTTATAAGG 23245

Qy         649 TTACAAGAATATTAGTTGGGCCTTATTTGTTTGTTTGTTGTTAATAATGTACATT 708
               ||||| |||||||| |||| |||||||| ||| ||| |||
Db       23246 TGACGCGTTTGTTGTTGATTGCGTGGTGTTTTTTGCGGTTCAGCATTATTGCCAACAATGTGCATT 23305

Qy         709 ATGCAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTTGCATTAAAGAAATTA 768
               | |||||  ||||  |||  |||| ||||||  ||||| |||  ||| || || |||||
```

```
Db         88017 TGACGCGTTGTTGTTGATTGGGTGTTTTTGCGTTCAGCATTATTGCCAACAATGTGCATT 88076
Qy           709 ATGCAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTTGCATTAAAGAAATTA 768
Db         88077 ACGCGGTTTATCAAAGCTGATGACGGGCATCAATTATTGGCTGATGCTGAAAGAGGTTA 88136
Qy           769 ATGAGATAACACAAATGCTGGCTTAACAATGATAGATAAATTCATATATCCATTGTTATTTG 828
Db         88137 CCGAAGTCGCAGCGCGGCGTGCGTCGATGTTGGATAAGTTGTGGCGTGCCTGTGTTGTGGG 88196
Qy           829 GTTTATTTGAAGTTGCTGTGTTTTTAAGTTTCATAAAAGAAAAGTATATAAAC 888
Db         88197 GCGTGTTGGAAGTCATGTTGTTTGCGCCTTGCCAGCTTGCCCGTAGACGCGCAT---- 88252
Qy           889 TTTCTTGGATTTTGACTTTATTTTTATGCTGTGATGATGTGTTTTTGTTCGAGCGT 948
Db         88253 ------TTTTCTGCCGATATACGTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCCGT 88304
Qy           949 ATACAACAAAATCCCATGAGCGTTTTATTTCACCTAACACTGTTTATTCTCGATTAAAAT 1008
Db         88305 CGTTCGACACGAAACAAGAGCACGGTATTTCGCCCAAACCGACATACAGCCGCATCAAAG 88364
Qy          1009 CCAATTATTTATCGTGTGGGTATTTTATAGGACGAATTGTTCCTTATGAGATATTTTCT 1068
Db         88365 CCAATTATTTCAGCTTCGGTTATTTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATT 88424
Qy          1069 TATCTAAATATTCCTCTTTATCATAAATCTAAGCCTATGAAATCGGGGCTCTCCGVAAAATT 1128
Db         88425 TAAGCAGGATTCCCGCCTTTAAGCAGCCTTGCTCCAAGCAAAATCGGG---CAGGGCAGTG 88481
Qy          1129 MARGAATATAATTTTAATTAATAAATGGGGGAAAGTGCGACCTCAAGTCATTTTAGTGCTTTTG 1188
Db         88482 TTCAAAAATCGTCCTGATTTATGGGCGAAAGCGGAGCGCGGGCCATTTGAAGCTGTTTG 88541
Qy          1189 GTTACGGGGAGAAAAACATCTCCTTTTTAGATAG---CTTAAAATATAAATCAGGAGCTC 1245
Db         88542 GCTACGGACGCGGAAACTTCGCCCGTTTTTACCCGCTGTCGCAAGCCGATTTTAAGCCGA 88601
Qy          1246 TTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTCTTTACCAATGTTTTTA 1305
Db         88602 TTGTGAACAAAGTTATTCCGCAGGCTTTATGACTGCAGTGTCCCGCCCAGTTTTTTCA 88661
Qy          1306 ATGCAATYCCTTAYCCAAATGGAATACAACAACAGATAGCTAAAGGAGATACGAATTTATTTA 1365
Db         88662 ATCGGATACCGCACGCCACGCGGCTTGGAACAAATCAGCGGCGGCGGATACCAAATATGTTCC 88721
Qy          1366 ATTTAGCGAAAGAGAGCAAGGCTTTCAGACATATTTTATTCAGCTGGAGGATGATA 1425
Db         88722 GCCTCGCCAAAGAGCAGGCGTATTTTACAGCGCGGAGGCGGAAAACGAGA 88781
Qy          1426 TGCATATGATCAATTTTTAGGAGGAGCTTGGATTGATGATATTCGTTTTCCAGATAATG 1485
Db         88782 TGGGGATTTTGAACTTAATCGGTAAGAAAATGGATAGACCCATCTGATTCAGCCGACGCAAC 88841
Qy          1486 AAGGGTATTCTTTAAGAGATTCAATGCCCGATAATAAAATTACTTCCTGCTTTAAAAAATA 1545
Db         88842 TTGGCTACGGCAACGGCGACAAATAGCCCGATGAGGAAGCTGCTGCGGTTGTTCGACAAAA 88901
Qy          1546 TTAATTTAGTAGATGGTTATCATTTTGTGTTTTGTTTTACATCATAGAGGGAGTCATATTCCCT 1605
Db         88902 TCAATTTGCAGCAGGGCAAGCATTTTATCGTGTTGCACCAACGCGGTTCGCACGCCCCAT 88961
Qy          1606 ATGGGGCATTATT 1618
Db         88962 ACGGCGCATTGTT 88974
```

```
RESULT 6
AAF21613
ID   AAF21613 standard; DNA; 172325 BP.
XX
AC   AAF21613;
```

KW Neisseria meningitidis; Neisseria gonorrheae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS Neisseria meningitidis.
XX
PN WO200066791-A1.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US005928.
XX
PR 30-APR-1999;   99US-0132068P.
PR 08-OCT-1999;   99WO-US023573.
PR 28-FEB-2000;   2000GB-00004695.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections.
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of Neisseria
CC meningitidis B (NMB). The sequences in AAF21607 to AAF21613
CC represent fragments of the NMB genomic sequence, as the sequence which
CC long to go in a record on its own it was split into 8 sequences which
CC overlap each other at the beginning and end of each sequence by 49980 bp
CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins
CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
CC primers which are used in the exemplification of the present invention.
CC The NMB genome and fragments from it have antibacterial activity, and can
CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
CC and/or antibodies which binds to the proteins can be used in compositions
CC for treating or preventing infection due to Neisserial bacteria or as a
CC diagnostic reagent for detecting the presence of Neisserial bacteria or
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
CC computer storage medium or computer databases can be used in a search to
CC identify open reading frames (ORFs) or coding sequences within the NMB
CC genome. The DNA sequences provide further opportunities to find antigenic
CC or immunogenic proteins which are more effective in vaccines than the
CC outer membrane proteins currently used
XX
SQ Sequence 349980 BP; 86473 A; 95646 C; 85908 G; 81953 T; 0 U; 0 Other;

Query Match   11.8%;   Score 196.2; DB 3; Length 349980;
Best Local Similarity 51.5%; Pred. No. 7.1e-26;
Matches 532;   Conservative 2;   Mismatches 481;   Indels 18;   Gaps   3;

Qy   589 CTGAAAGTTATTTGTCTGTTTGTTTTGTTTTGTTTTTGTTTTTTTTGTTTTTGTTTAATATAGAA 648
            ||||||||||||||||||||||| |||||||| ||||||||||||||||||||||||||||
Db   323186 CGGAAACGTTTGCCTGACATTTGGCGTGTTTTTTGCGTTAATACTCACTTGCATTTAAAGAAATTA 323245

Qy   649 TTACAAGAATATAGTTGGCGCTTTATTTGCAATAAGTATGTTGTTGTTAATGTACATT 708
            |||||||||||||||||||||||||||
Db   323246 TGACCGCGTTTGTTGTTGATTGGCGTGTTTTTTGCGTTCAGCATTATTGCCAACAATGTCATT 323305

Qy   709 ATGGCAGTATACCAATCTTCGGATTGAATGATGACGGCATCAATTATTGGCTGATGTCGAAAGAAATTA 768
            ||||||||||||||
Db   323306 ACGCGGTTTATCAAAGCTCGATGATGACCATCAATTATTGGCTGATGTCGAAGAGGTTA 323365

Qy   769 ATGAGATAACAAATGCTGGCTTAACAATGATAGATAAATTCATTATATCCATTGTTATTTG 828
            |||||||||||||||||||||||||
Db   323366 CCGAAGTCGGCAGCGCGGGTGCGTCGATGTTGGATAAGTTGTGGCTGCCTGGTTGTGTGGG 323425

Db   23306 ACGCGGTTTATCAAAGCTGGATGACGGGCATCAATTATTGGCTGATGTCGAAAGAGGTTA 23365
            |||||||||||||||||||||||||
Qy   769 ATGAGATAACAAATGCTGGCTTAACAATGATAGATAAATTCATTATATCCATTGTTATTTG 828

Db   23366 CCGAAGTCGGCAGCGCGGGTGCGTCGATGTTGGATAAGTTGTGGCTGCCTGGTTGTGTGGG 23425

Qy   829 GTTTATTTGAAGTTGCTGTGTTTTTAAGTTTAAGTTCATAAAAAGAAAAGTATATAAAC 888
            | |||||||||||
Db   23426 GCGTGTTGGAAGTCATGTTGTTTTTGCAGCCTTGGCCAAGTTCCGCCGTAAGACGGCAT---- 23481

Qy   889 TTTCTTGGATTTTTATTTTTTATGCTGTGATGATGTATGTTTTGTTCGAGCGT 948
            |||||||||
Db   23482 --------TTTTCTGCGATATACTGTTTGCCTTCCTATATGCTGATGTTTCTGTGCCTT 23533

Qy   949 ATACAACAAAATCCCATGACGCGTTTTATTCACCTAACACTGTTTATTCTCGATTAAAAT 1008
            ||||||
Db   23534 CGTTCGACACGCGAAACAAGACACGTATTTCGCCCAAACCGACATACAGCCCGCATCCAAAG 23593

Qy   1009 CCAATTATTTATCCGTTGGGTTATTTTATAGGACCGAATTGTTCCTTATGAGATATTTCTT 1068
            ||||||||||||||||||||| ||||||||||
Db   23594 CCAATTATTTCAGCTTCCGGTTTATTTTGTCGGACGCGTGTTGCCGTATCCAGTTGTTTGATT 23653

Qy   1069 TATCTAATATTCCTCTTTATCATAAATCTAAGCCTAGAAATCGGGCTCTCCGVAAAATT 1128
            ||||||||||||||||||
Db   23654 TAAGCAGGATTCCCGCCCTTTAAGCAGCCTGCTCCAAGCAAAATCGGG---CAGGGCAGTG 23710

Qy   1129 MARGAATATAATTTTAATTAATGGGGAAAGTGCGACCTCAAGTCATTTTAGTGCTTTCTT 1188
            |||||||||||||||||||||||||||
Db   23711 TTCAAAAATATCGTCCTGATTATGGGCGAAAGCGAAAGCCATTTGAAGCTGTTTG 23770

Qy   1189 GTTACGGGGAGAAAAACATCTCCTTTTTAGATAG---CTTAAAATATAAATCAGGAGCTC 1245
            |||||||||||||| |||||||
Db   23771 GCTACGGACGCGGAAACTTCGCCGTTTTTAACCCGGCTGTCGCAGCCGATTTTAAGCCGA 23830

Qy   1246 TTGTGGTAAAACTTATTCAGGAGGAAAGCTAACAGACAATTTCTTTACCAATGTTTTTTA 1305
            |||||
Db   23831 TTGTGAAACAAAGTTATTCCGCGACGGCCTTATGACTGCAGTGTCCCTGCCCAGTTTTTCA 23890

Qy   1306 ATGCAAATYCCTTAYCCAAATGGAATAACACAGTAGCTAAGGAGATACGAATTTATTTA 1365
            |||||||||||||||
Db   23891 ATCGGATACCGGACCACGCCCAACGGCTTGGAACAAATCAGCGGCGGCGGATACCAATATCGTTCC 23950

Qy   1366 ATTTAGCGGAAAGAGCAAGGGCTTTCAGACATATTTTATTCAGCTAGGGGATGATA 1425
            |||||||
Db   23951 GCCTCCGCCAAAGAGCAGGGCTATGAAACGTATTTTTTACAGCGCGCAGGCGGAAAACGAGA 24010

Qy   1426 TGCATAATGATCAATTTTTAGGAGGAGCTTGGATTGATGATGATTGTTTCCAGATAATG 1485
            ||||||||
Db   24011 TGGCGATTTGAACTTAATCGGTAAGAAATGGATAGACCATCTGATTCAGCCGACGCAAC 24070

Qy   1486 AAGGGTATTCTTTAAGAGATTCAAATGCCGATAATAAATTACTTCCCTGCTTTTAAAAATA 1545
            |||
Db   24071 TTGGCTACGCAACGCGGCAACATGCCCGATGAGAAGCTGCTGCCGTTGTTCGACACAAAA 24130

Qy   1546 TTAATTTAGATAATGGTTATCATTTGTTGTTTACATCATAGAGGGGAGTCATATTCCCT 1605
            | ||||||||||||||
Db   24131 TCAATTTGCAGCAGGGCAAGCATTTTATCGTGTTGCACCAACGCGGTTCGCACGCCCAT 24190

Qy   1606 ATGGGGCATTATT 1618
            |||||||||||||
Db   24191 ACGGCGCATTGTT 24203

RESULT 7
AAF21612
ID AAF21612 standard; DNA; 349980 BP.
XX
AC AAF21612;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:113.
XX

```
Qy    829 GTTTATTTGAAGTTGCTGTGTGTTTTTAAGTTAAGTTCATAAAAAGAGAAAAGTATATAAAC 888
          |||||||                    ||||     ||| ||  |    |||
Db 323426 GCGTGTTGGAAGTCAGTTCATGTTGTTTTGCAGCCTTGCCCAAGTTCCGCCGTAAGACGGCAT--- 323481

Qy    889 TTTCTTGGATTTTGACTTTATTTTTATTTTATGTCTGTGATGATGTATGTTTTTGTTCGAGCGT 948
          |||||||                  |  |  ||||  |||| |||||
Db 323482 ------TTTTCTGCCGATATACTGTTTGCCTTCCTAATGCTGATGATGATTTTCGTGCGTT 323533

Qy    949 ATACAACAAAATCCCATGAGACGTTTTATTTCACCTAACACTGTTTATTCTCTGATTAAAAT 1008
          |||||||    |||||| |   |||||     ||  |||| |   ||  || |
Db 323534 CGTTCGACACGAAACAAGACACGTATTTCGCCCAAACGACATACAGCCGCATCAAAG 323593

Qy   1009 CCAATTATTTATTTCGTTGGGTTATTTATTTATAGGACGAATTGTTCCTTATGAGATATTTTCTT 1068
          |||||||||             |||  ||| ||   |    |||||| |  |
Db 323594 CCAATTATTTCAGCTTCGGTTATTTGTCGGACGCGTGTTGCCGTATCAGTGTTTGATT 323653

Qy   1069 TATCTAATATTCCTTTTATCATAAATCTAAGCCTATGAAATCGGGCTCTCCGTAAAATT 1128
          ||||||||||       |  ||  |   ||  ||||   |  | ||  :  |
Db 323654 TAAGCAGGATTCCCGCCTTTAAGCAGCAGCGTTGGGAAATCGGG---CAGGGCAGTG 323710

Qy   1129 MARGAATATAATTTAATTAAATGGGGAAAGTGCGACCTCCAAGTCATTTTAGTGTTCTTTC 1188
          |||||||    |||  || |||||||  |||  |||  |||  ||||||||||  |
Db 323711 TTCAAAATATCGTCCTGATTATGGCGAAAGCGCGGGCCATTTGAAGCTGTTTG 323770

Qy   1189 GTTACGGGGAGAAAAAACATCTCCTTTTTTTAGATAG---CTTAAATATAAATCAGGAGCTC 1245
          ||||||          ||||||||     |||||| |       |  || || |||||||
Db 323771 GCTACGGACGCGAAACTTCGCCGTTTTTAACCCGCTGTGTCGCAAGCCGATTTTAAGCCGA 323830

Qy   1246 TGGTGGTAAAACTTATTCAGGAGGAAGCTAACAGCAATTTCTTACCAATGTTTTTA 1305
          ||||                  ||| ||| ||   ||||||||| | ||||||  |
Db 323831 TTGTGAAACAAAGTTATTCCGCAGGCTTTTATGACTGCAGTGTCCCTGCCCAGTTTTTCA 323890

Qy   1306 ATGCCAATYCCTTAYCCAAATGGAATACAACAGATAGTAAAGAGAATACGAATTTATTTA 1365
          ||||           : || |||||||    ||   |  ||
Db 323891 ATGCGGATACCGCAAGCGGCTTGGACAAATCAGCCGGCGGATACCAATATGTTCC 323950

Qy   1366 ATTTAGCGAAAGAGCAAGGCGTTCAGAGAGAGCAAGGCTTTCAGCTTCAAGCTCAAGCTAGGGATGATA 1425
          |||||||||||          |  ||| ||||||| | :|  |||  |    |
Db 323951 GCCTCGCCAAAGAGCAGGGGCTATGAAACGTATTTTTACAGCGGCGGAAAACGAGA 324010

Qy   1426 TGCATATGATCAATTTTTAGGAGGAGCTTGGATTGGATGATGATATTCGTTTCCAGATAATG 1485
          |||   |   |||| | |   |||| | |||   ||||| | |  |    |  |||
Db 324011 TGGCGGATTTTGACTTAATTCGGTAAGAAATGGATGTAAATTCAGCCGACGCAAC 324070

Qy   1486 AAGGGTATTCTTTAAGAGATTCAATGCCTGATAATAATAAATTACTTCCTGCTTTTAAAATA 1545
          ||| ||||| ||||     | |||  |  |||  | ||||   | |||| | |
Db 324071 TTGGCTACGGCACACGGGCAACAATATGCCCGATGAGAGCGTGCTGCCGGTTGTTCGCACAAA 324130

Qy   1546 TTAATTTAGATAATAATGGTTATCATTTTGTTGTTTACATCATAGAGGGAGTCATCATTCCCT 1605
          |||||||||||             ||||||| | ||||  || | |    |||
Db 324131 TCAATTTGCAGCAGGGCAAGCATTTTATCGTGTTCAAATGCGTGTTCGCACCGTGGCAT 324190

Qy   1606 ATGGGGCATTATT 1618
          ||||||| ||||
Db 324191 ACGGTGCGCATTGT 324203
```

RESULT 8
ABZ41628
ID   ABZ41628 standard; DNA; 1467 BP.
XX
AC   ABZ41628;
XX
DT   07-MAR-2003 (first entry)
XX
DE   N. gonorrhoeae nucleotide sequence SEQ ID 7845.
XX
KW   Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX
OS   Neisseria gonorrhoeae.
XX
PN   WO200279243-A2.

CC   The present invention relates to proteins from Neisseria gonorrhoeae.
CC   Also disclosed are the nucleic acid molecules encoding the proteins and
CC   antibodies that specifically bind to the proteins. The composition
CC   comprising the protein, nucleic acid or antibody is useful for the
CC   manufacture of a medicament for treating or preventing N. gonorrhoeae
CC   infection, this may be in the form of a vaccine or gene therapy.
CC   Sequences given in records AB237706-AB242016 represent nucleic acid
CC   molecules of the invention
XX
SQ   Sequence 1467 BP; 356 A; 358 C; 377 G; 376 T; 0 U; 0 Other;

```
Query Match             11.2%;  Score 185.8;  DB 10;  Length 1467;
Best Local Similarity   50.7%;  Pred. No. 3.4e-24;
Matches  524;  Conservative    4;  Mismatches  487;  Indels   18;  Gaps    3;

Qy  589 CTGAAAGTTATTTGTCTGTTGTTGTTGTTTGTTGTTTTTGTTTTTTCAAAATATAGAA 648
         ||  |  | | ||||  |  ||||  |  |||| ||| || |||  |  ||  |||
Db    5 CGGAAACGTTGCGCTGACATTTATGATTGCTGCGCGTATCTGTTTGCGCGTTATAAGG 64

Qy  649 TTACAAGAATATTAGTTGGCGCTTTATTTCGTAATAAGTATGTTGTTAATATGTACATT 708
         || ||| |||| |   |||  |  | ||| |   |||| |  |  || |   |||||
Db   65 CTTCGCGGCTGCTATATGCGCGTGTTTTTCCGCGGTTCAGCCATGCCAACAATGTGCATT 124

Qy  709 ATGCAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTTGCATTAAAGAAAATTA 768
         || ||| |  ||  ||| || | | |  | |||  |  | | ||| || || |  |
Db  125 ACGCCGTTATCAAAGCTGATGACGGGTATTAACTATTGGCTGCGATGCTGAAAGAGGTTA 184

Qy  769 ATGAGATAACAAATGCTGGCTTAACAATGATAAATTCATATCATCCATTGTTATTTG 828
         || ||| |  || |  | |||||  |    || ||| |   ||  |  ||||  || |||
Db  185 CCGAAGTCGGCAGGTCGCGGCGGCGCGTCGATAAGTGTTGGATAAGTTCATGAGCCTTGTTATTTG 244

Qy  829 GTTTATTTGAAGTTGCTGTGTGTTTTTAAGTTAAGTTCATAAAAAGAGAAAAGTATATAAAC 888
         |||||   | |||  |||| | | ||  |  |||   | |||  || |  ||||
Db  245 GCGTGGCGGAAGTCAGTCATGTTGTTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCAT---- 300

Qy  889 TTTCTTGGATTTTTGACTTTTTATTTTTATGCTGTGTTGCCTTCCTAATGCTGATGATGAGTCG 948
         |||||    |||||||     |||  |||  |  | |||||
Db  301 ------TTTTCTGCCGATATACTGTTTGCCTTCCTAATGCTGATGATGATTTTCGTGCGTT 352

Qy  949 ATACAACAAAATCCCATGAGACGTTTTATTCACCTAACACTGTTTATTCTCTGATTAAAAT 1008
         ||||||  |||||||| |  | |||||  | || | || |    || || ||| |||
Db  353 CGTTCGACACGAAACAAGACACGTATTTCGCGGGCGCGTGTTCCTTATGAGATATTTTCTT 412

Qy 1009 CCAATTATTTATCGTTGGTTATTTATTTATAGGACGAATTGTTCCGTATCAGTGTTTGATT 1068
         ||||||||          |||||  ||| ||   |     ||||| |  |
Db  413 CCAATTATTCAGCTTCGGTTATTTTGTTGTTGCCGTGTTATCAGTGTTTGCCGTTTGATT 472

Qy 1069 TATCTAATATTCCTCTTTATCATAAATCTAAGCCTATGAAATCGGGCTCTCCGVAAAATT 1128
         ||| |||||       |  || |   ||   ||   |  | ||| |  |;  |
Db  473 TAAGCAAGATCCCGTCTGTCAAACAGCCTGCTCCAAACAGCCAGGCAGGCAGTATTC 532

Qy 1129 MARGAATATAATTTAATTAAATGGGGAAAGTGCGACCTCCAAGTCATTTTAGTGTCTTTG 1188
         ||||||||    |||  || ||||||  |||  |||  |||  ||||||||||  |
Db  533 AAAATATCGTCCTG---ATTATGGCGAAAGCGCGAAAGCGGCCATTTGAAATTGTTTG 589
```

PS   Claim 9; Page 208; 524pp; English.
XX
CC   Nucleotide sequences AAZ11972-Z12358 represent open reading frames (ORFs)
CC   of Neisseria meningitidis and N. gonorrhoeae which encode antigenic
CC   proteins (see AAY38499-Y38944). The antigenic proteins, their fragments,
CC   their nucleic acids and antibodies are used for diagnosis, prevention (as
CC   vaccines) or treatment of Neisseria infections, such as meningitis,
CC   septicaemia and gonorrhea. Both organisms are closely related. Fragments
CC   of the nucleic acids are useful as hybridisation probes and antisense
CC   reagents
XX
SQ   Sequence 1575 BP; 385 A; 387 C; 392 G; 411 T; 0 U; 0 Other;

     Query Match   11.2%; Score 185.8; DB 2; Length 1575;
     Best Local Similarity   50.7%; Pred. No. 3.4e-24;
     Matches 524; Conservative   4; Mismatches 487; Indels   18; Gaps   3;

Qy   589 CTGAAAGTTATTGTCTGTTGTTGTTTGTTGTTTGTTTTTGTTTTTTCAAAATATAGAA 648
          ||||| ||||||||||||||||||| |||| ||||||||||||||||||||
Db   110 CGGAAAACGTTTGCGCTGACATTTATGCTGCGCGCTGTACGTTCGCGGTTATAAGG 169

Qy   649 TTACAAGAATATTAGTTGGCGCTTTATTTGCAATAAGTATTGTGTTAATAATGTACATT 708
          |||||| ||||||||||
Db   170 CTTCGGGGCTGCTGAATTGCGGTGTTTTTTCGCGGTTCAGCATGATTGCCAACAATGTGCATT 229

Qy   709 ATGCAGTATACCAATCTGATTGGACCTGTTAATTACTCACTTGCATTTAAAGAAATTA 768
          ||||||||||||||||||||||
Db   230 ACGCGGTTTATCAAAGCTGGATGACGGGTATTAACTATTGGCTGATGCTGAAAGAGGTTA 289

Qy   769 ATGAGATAACAAATGCTGGCTTAACAATGATAGATAAAAATTCATATATCCATTGTTATTTG 828
          |||||||||||||||||||||||
Db   290 CCGAAGTCGGCAGCGGCGGCGCGTCGATGTTGGATAAGTTGTGGCGTCGCCTGCTTTGTGGG 349

Qy   829 GTTTATTTGAAGTGCTGCTGTGTTTTTAAGTTTAAGTTTCATAAAAGAAAAGTATATAAAC 888
          ||||||||||||||||
Db   350 GCGTGGCGGAAGTCATGTTGTTTGTTGCAGCCTTGCCAAGTTCCGCCAAGTCCGTAAGACGCAT--- 405

Qy   889 TTTCTTGGATTTTGACTTATTTATTTTATGCTGTGATGAGTATGTTTTTGTTCGAGCGT 948
          ||||||||||||||
Db   406 --------TTTTCTGCGATATACTGTTTGCTTCCTAATGCTGATGATTTTCGTGCGTT 457

Qy   949 ATACAACAAATCCCATGAGCGTTTATTTCACCTAACACTGTTTATTCTCGATTAAAAT 1008
          ||||||||||||
Db   458 CGTTCGACACGCAAACAAGAGCACACGGTATTTCGCCCAAACCGACATACAGCCGCATCCAAAG 517

Qy   1009 CCAATTATTATCGTTCGGTTATTTGTCGGTTATTTGTCCGTGTTGCCGTATCAGTTGTGTTGATT 1068
          ||||||||||||||||||||||||||||
Db   518 CCAATTATTATCGTTCGGTTATTTGTCGGTTATTTGTCCGTATCAGTTCAGTTGTGTTGATT 577

Qy   1069 TATCTAATATTCCTCTTTATCATAAATCTAAGCTATGAAATCGGCGCTCTCCGVAAAATT 1128
          ||||||||
Db   578 TAAGCAAGATCCCTGTGTTCAAACAGCCTGCTCCAAGCAAAATCGGGCAAGGCAGTATTC 637

Qy   1129 MARGAATATAATTTTAATTAATGGGGAAGTGCGACCTCAAGTCATTTAGTGCTTTTG 1188
          ||:| ||||| ||||
Db   638 AAAATATCGTCCTG---ATTATGGGCGAAAGCGAAAGCGCGCGCATTTGAAATTGTTTG 694

Qy   1189 GTTACGGGAGAAAAACATCTCCTTTTTTAGATAG---CTTAAAATATAAATCAGGAGCTC 1245
          ||||||| || ||
Db   695 GTTACGGGCGGCGAAACTTCGCCGTTTTAACCCGGCTGTCGCAAGCCGA 754

Qy   1246 TTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTTCTTTACCAATGTTTTTA 1305
          |||| ||| ||||| |||
Db   755 TTGTGAAACAAGTTATTCCCAGGCTTTATGACGGCCAGTATTCCCTGCCCAGTTTCTTTA 814

Qy   1306 ATGCAATYCCTTAYCCAAATGGAATACACAGATAGTCAAGGAGATACGAATTATTTA 1365
          ||||| |:
Db   815 ACGTCATACCGCACGGCCACGGCCTTGGAACAAATCAGCGGCGGCGGATACCAATATGTTCC 874

Qy   1366 ATTTAGGCAAAGAGCAAGGCTTTCAGACACATATTTTATTCAGCTCAAGCTAGGGATGATA 1425
          ||||||||||||| |||||||| |||||
Db   875 GCCTCCGCCAAAGACGAGGGCTATGAAACGTATTTTTACAGTGCCCAGCTGAAAACCAAA 934

Qy   1426 TGCATATGATCAATTTTTTAGGAGGAGCTTGGATTGATTGATATTCGTTTCGTTTTCCAGATAATG 1485
          ||||||||||
Db   830 TGGCAATTTTGAACTTAATCGGTAAGAAATGATAGACCATCTGATTCAGCCGACGCAAC 889

Qy   1486 AAGGGTATTCTTTAAGAGATTCAATGCCTGATAATAAATTACTTCCTGCTTTTAAAAATA 1545
          ||| ||
Db   890 TTGGCTACGGCAACGGCGACAATATGCCCGATGAGAAGTCGTGCTGCCTTGTTGTTCGACAAAA 949

Qy   1546 TTAATTTAGATAATGGTTATCATTTTGTGTTTACATCATAGAGGGAGTCATATTCCCT 1605
          |: ||
Db   950 TCAATTTCAGCAGGCAGGCAGGCATTTTATCGTGTTGCACCAACGCGGTTCGCACGCCCCAT 1009

Qy   1606 ATGGGGCATTATT 1618
          |||| ||||
Db   1010 ACGGCGCATTGTT 1022

RESULT 9
AAZ12107
ID   AAZ12107 standard; DNA; 1575 BP.
XX
AC   AAZ12107;
XX
DT   08-OCT-1999 (first entry)
XX
DE   Neisseria gonorrhoeae complete ORF81 sequence.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS   Neisseria gonorrhoeae.
XX
PN   WO924578-A2.
XX
PD   20-MAY-1999.
XX
PF   09-OCT-1998; 98WO-IB001665.
XX
PR   06-NOV-1997; 97GB-00023516.
PR   14-NOV-1997; 97GB-00024190.
PR   18-NOV-1997; 97GB-00024386.
PR   27-NOV-1997; 97GB-00025158.
PR   10-DEC-1997; 97GB-00026147.
PR   14-JAN-1998; 98GB-00000759.
PR   01-SEP-1998; 98GB-00019016.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
XX
DR   WPI; 1999-327407/27.
DR   P-PSDB; AAY38653.
XX
PT   Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT   diagnosis, treatment and prevention of infection.
XX

```
Qy    589  CTGAAAGTTATTTGTCTGTTGTTGTTGTTGTTGTTTTTCAAAATATAGAA  648
                ||||   |||||||||||| |||||||||||||  || |||||||
Db    110  CGGAAACGTTTGCGCTGACATTTATGATTGCTGCGGCGTATCTGTTGTTGCGGCGTTATAAGG  169

Qy    649  TTACAAGAATATTAGTTGGGCGCTTATTTGTTTATTTGCAATAAGTAGTATTGTGTTAATAATGTACATT  708
                  |||   |||||||||||||                   ||||||||||||||||
Db    170  CTTCGCGGCGTGCTGATTGCGCGTGTTTCGCGGTTTTTCGCGTTCAGCATGATTGCCAACAATGTCATT  229

Qy    709  ATGCAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTTGCATTTAAAGAAATTA  768
                 |||                |||  ||||| || |||||||||||||
Db    230  ACGCGGTTTATCAAGCTGGATGACGCGGTATTAACTGGCTGATGCTGAAAGAGGTTA  289

Qy    769  ATGAGATAACAAATGCTGGCTTAACAATGATAGATAAATTCATATATCCATTGTTATTTG  828
                ||||||                   ||||  |||| |||||||||||||||||||
Db    290  CCGAAGTCGGCAGCCGGCGCGTGCATGATGTTGGATAAGTTGTGGCGCCTGCTTGTTGTGGG  349

Qy    829  GTTTATTTGAAGTTGCTGTGTTTTGTGTTTTAAGTTTAAGTTTCATAAAAAGAAAAGTATATAAAC----  888
                ||||| || ||||| || |||||||||||||||     ||| ||||||||||||||||
Db    350  GCGTTGGCGGAAGTCATGTTGTTTTGCGGCCTTGCGATGTATGTTTTGTTTTGTTCGAGCGT  405

Qy    889  TTTCTTGGATTTTTGACTTATTTATTTTATGCGTGTGTGATGATGATTTTCGTGCGT  948
                       ||||     ||| |||  |||||||  |||||||||||||||||
Db    406  ------TTTCTGCCGATATACTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTT  457

Qy    949  ATACAACAAAATCCCATGAGCGTTTTATTCACCTGAACACTGTTTATTCTCGATTAAAAT  1008
                ||| ||||||||||||      |||||||||||||||||||||||||||||
Db    458  CGTTGCACGCAAACAAGAGCACGGTATTTCGCCCAAACCGACACATACAGCCGCATCAAAG  517

Qy    1009  CCAATTATTTATTATCGTTGGGTTATTTTATTAGGACGAATTGTTCCTTATGAGATATTTCTT  1068
                  |||||||| |||    ||| || || |||||||||||||||||||||||
Db    518  CCAATTATTTCAGCTTCGGTTATTTGTCGGGCGCGTGTTGCCGTATCAGTTGTTTGATT  577

Qy    1069  TATCTAATATATTCCCTTCTTATCATAAATCTAAGCCTATGAAATCGGGCTCTCCGVAAAATT  1128
                :  ||  |||||||  ||||||||| |||||||||||||||||||||||| |
Db    578  TAAGCAAGATCCCCTGTTCAAACAGCCTGCTCCAAGCAAATCGGGCAACAGCAGTATTC  637

Qy    1129  MARGAATATAATTTTAATTAAAtGGGGGAAAGTGCGACCTCAAGTCATTTTAGTGCTTTTG  1188
                : :|     |||| |||||||||||||||||||       ||||||||||||||||
Db    638  AAAATATCGTCCTG---ATTATGGGCGAAAGCGAAAGCGGCGGCGATTTGAAATTGTTTG  694

Qy    1189  GTTACGGGGAGAAAAACATCTCCTTTTTAGATAG---CTTAAAATATAAAATCAGGAGCTC  1245
                |||||||||      ||| || | ||||||||       |||  ||||||| |||||||
Db    695  GTTACGGGGCGGAAACTTCGCCGTTTTAACCCGGCTGCGTGCCAAGCCGATTTTAAGCCGA  754

Qy    1246  TTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTTCTTTACCAATGTTTTTTA  1305
                ||| ||||  |||||||||||  ||||||| || |||||||||| |||  ||| |||||
Db    755  TTGTGAAAACAAAGTTATTCCGGCAGGCTTTATGACGGCAGTATCCCTGCCCAGTTTCTTTA  814

Qy    1306  ATGCAATYCCTTAYCCAATGGAATACAACAGATAGTCTAAAGACGCGGCGGGCGGATAACCAATATGTTCC  1365
                |||||     ||    ||||| || ||| |:|| |||||||||| ||||||| |:|||||||
Db    815  ACGTCATACCGCACGCCACGCGTTGGAACAAATCAGCCGGCGGGATATCCATATATGTTCC  874

Qy    1366  ATTTAGCGAAAGAGCAAGGCTTTCAGACACATATTTTTATTCAGTCAAGCTAGGGATGATA  1425
                 ||||||| |||||||||||||||||||| |||  ||||| |||||| ||||  |||||
Db    875  GCCTCGCCAAAGAGCACGCAGGGCTATGAAACGTATTTTTACAGTGCCCAGGCTGAAAACCAAA  934

Qy    1426  TGCATATGATCAATTTTCCAGATAATG  1485
                |||  ||||||||   |||||  |||||||||||||||||||||| |||
Db    935  TGGCAATTTTGAACTTAATCGGTAAGAAATGGATAGACCATCTGATTCAGCCGACGCAAC  994

Qy    1486  AAGGGTATTCTTTAAGAGATTCAATGCCTGATAATAAAATTACTTCCTGCTTTTAAAAATA  1545
                ||| ||| |||  |||  ||||||||  |||||||||||||||||||||||||
Db    995  TTGGCTACGGCAACGGCGACAATATGCCCGATGAGAAGCTGCTGCCGTTGTTGTTGACAAAA  1054

Qy    1546  TTAATTTAGTAAATGGTTATCATTTTGTGTTTTACATCATAGAGGAGTCATATTCCCT  1605
                 |||||  ||||| |||||||||||||| |||||||| ||||||||||||||||||
Db    1055  TCAATTTGCAGCAGGCAGGCATTTTATCGTGTTGCACCAACGCGGTTCGCACGCCCCAT  1114

Qy    1606  ATGGGGCATTATT  1618
                |||| |||||||
Db    1115  ACGGCGCATTGTT  1127
```

RESULT 10

```
AAA15323
ID    AAA15323 standard; DNA; 1575 BP.
XX
AC    AAA15323;
XX
DT    15-SEP-2003  (revised)
DT    04-SEP-2000  (first entry)
XX
DE    DNA encoding a polypeptide of a Neisseria pathogenic strain.
XX
KW    Pathogenic strain; Neisseria; vaccine; Neisseria infection; ss.
XX
OS    Neisseria gonorrhoeae.
XX
FH    Key              Location/Qualifiers
FT    CDS              1..1575
FT                     /*tag= a
XX
PN    WO200026375-A2.
XX
PD    11-MAY-2000.
XX
PF    28-OCT-1999;  99WO-FR002643.
XX
PR    30-OCT-1998;  98FR-00013693.
XX
PA    (INMR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
PA    (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI    Aujame L,  Bouchardon A,  Renauld-Mongenie G,  Rokbi B,  Nassif X;
PI    Tinsley C,  Perrin A;
XX
DR    WPI; 2000-365622/31.
DR    P-PSDB; AAY93293.
XX
PT    New polypeptide specific for pathogenic Neisseria useful in therapeutic
PT    or preventative vaccines and for diagnosis.
XX
PS    Claim 4; Page 137-139; 187pp; French.
XX
CC    The present sequence encodes a protein that is specific for pathogenic
CC    strains of Neisseria. The polynucleotides, polypeptides, or their
CC    antigenic fragments, are used in vaccines to treat or protect against
CC    Neisseria infections, particularly by N. meningitidis. The polynucleotide
CC    sequence is also used for recombinant production of the polypeptide and
CC    to produce attenuated Neisseria strains that overexpress it, or express
CC    it in a non-toxic mutant form. (Updated on 15-SEP-2003 to standardise OS
CC    field)
XX
SQ    Sequence 1575 BP; 385 A; 387 C; 392 G; 411 T; 0 U; 0 Other;
```

```
Query Match             11.2%;  Score 185.8;  DB 3;  Length 1575;
Best Local Similarity   50.7%;  Pred. No. 3.4e-24;
Matches 524;  Conservative    4;  Mismatches 487;  Indels    18;  Gaps    3;
```

```
Db   350 GCGTGTTGGAAGTCATGTTGTTTGCAGCCTTGCCCAAGTTCCGCCGTAAGACGCAT----  405
         |   ||| |||| |||||
Qy   889 TTTCTTGGATTTTGACTTATTTTATGCTGTGATGATGTATGTTTTGTTCGAGCGT       948
         | |||||     |
Db   406 ------TTTTCTGCCGATATACTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTT    457
         | || | |||
Qy   949 ATACAACAAAATCCATGAGCGTTTATTCACCTAACACTGTTATTCTCGATTAAAAT     1008
                |||  |    ||||  |
Db   458 CGTTCGACACGAACAGGAACACGTGGTATTTCGCCCAACCGACATACAGCCGCATCAAAG   517
           ||||||||||||
Qy  1009 CCAATTATTATTCGTTGGGTTATTTTATAGGACGAATTGTTCCTTATGAGATATTTTCTT  1068
         |||||||||  |||  ||||||  |||
Db   518 CCAATTATTCAGCTTCGGTTATTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATT    577
           |  |  |    |     |||   |
Qy  1069 TATCTAATATTCCTTTTATCATAAATCTAAGCCTATGAAATCGGGCTCTCCGVAAAATT   1128
         |   |||    |   |  |
Db   578 TAAGCAAGATTCCTGTGTTCAAACAGCCTGCTCCAAGCAGAATCGGGCAAGGCAGTATTC   637
Qy  1129 MARGAATATAATTTAATTAATGGGGGAAAGTGCGACCTCAAGTCATTTTAGTGCTTTG    1188
         :|:  |                ||||| |||| |||||       ||||||
Db   638 AAAATATCGTCCTGA---TTATGGGCGAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTG  694
Qy  1189 GTTACGGGAGAAAAACATCTCCTTTTT---AGATAGCTTAAAATATAAATCAGGAGCTC   1245
         || |||| | ||| |     || ||||    |||||||      ||||||  |||||| |
Db   695 GCTACGGGCGGGCGAAACTTCGCCGTTTTGACCCAGCTTTGCCAAGCCGATTTTAAGCCGA  754
Qy  1246 TTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTTCTTTACCAATGTTTTTTA   1305
         ||||  |||||| |||| |||  |||||||||| ||  |  ||||| |
Db   755 TTGTGAAACAAAGTTATTCCCGCAGGCTTTATGACGGCAGTATCCCTGCCCCAGTTTCTTTA  814
Qy  1306 ATGCAATYCCTTAYCCAAATGGAATACAACAGATAGTCTAAAGGAGGATA            1353
         |  ||  ||  ||    ||  |||  ||| | ||||||||  ||||||
Db   815 ACGTCATACCGCATGCCAACGCTTGAACAAATCAGCGGCGGCGATA                862
```

RESULT 11
AAZ12106
ID   AAZ12106 standard; DNA; 1275 BP.
XX
AC   AAZ12106;
XX
DT   08-OCT-1999 (first entry)
XX
DE   Neisseria meningitidis strain A complete ORF81 sequence.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW   treatment; Neisseria infection; meningitis; septicaemia; gonorrhea; ss.
XX
OS   Neisseria meningitidis.
XX
PN   WO9924578-A2.
XX
PD   20-MAY-1999.
XX
PF   09-OCT-1998; 98WO-IB001665.
XX
PR   06-NOV-1997; 97GB-00023516.
PR   14-NOV-1997; 97GB-00024190.
PR   18-NOV-1997; 97GB-00024386.
PR   27-NOV-1997; 97GB-00025158.
PR   10-DEC-1997; 97GB-00026147.
PR   14-JAN-1998; 98GB-00000759.
PR   01-SEP-1998; 98GB-00019016.
XX
PA   (CHIR-) CHIRON SPA.
XX
PI   Masignani V, Rappuoli R, Pizza M, Scarlato V, Grandi G;
XX
DR   WPI; 1999-327407/27.
DR   P-PSDB; AAY38652.
XX
PT   Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT   diagnosis, treatment and prevention of infection.
XX
PS   Claim 9; Page 206; 524pp; English.
XX
CC   Nucleotide sequences AAZ11972-212358 represent open reading frames (ORFs)
CC   of Neisseria meningitidis and N. gonorrhoeae which encode antigenic
CC   proteins (see AAY38499-Y38944). The antigenic proteins, their fragments,
CC   their nucleic acids and antibodies are used for diagnosis, prevention (as
CC   vaccines) or treatment of Neisseria infections, such as meningitis,
CC   septicaemia and gonorrhea. Both organisms are closely related. Fragments
CC   of the nucleic acids are useful as hybridisation probes and antisense
CC   reagents
XX
SQ   Sequence 1275 BP; 300 A; 304 C; 319 G; 352 T; 0 U; 0 Other;

Query Match   7.3%; Score 120.8; DB 2; Length 1275;
Best Local Similarity 50.0%; Pred. No. 1.6e-12;
Matches   384; Conservative   4; Mismatches 362; Indels   18; Gaps   3;

```
Qy   589 CTGAAAGTTATTTGTCTGTGTTGTTGTTGTTGTTTTGTTTTTCAAAATATAGAA       648
         | |||| ||| |||| ||| || | || || |  ||| |  ||||||| ||||
Db   110 CAGAAACGTTTGCGCTGACATTTGTGATTGCTGCGCTGATCTGTTTGCGCGTTATAAGG   169
Qy   649 TTACAAGAATATTAGTTGGCGCCTTTATTTGCAATAAGTAATTGTTGTTAATAATGTACATT 708
         | ||  || ||| ||| | |     |||| |  |  ||| | |||||    || ||| ||||
Db   170 CAACGGCGTTTGTTGATTGCGCGGTGTTTTTCGCGTGATTATTGCGCCATTATTGCCAACACATGTGCATT 229
Qy   709 ATGACAGTACCAAATCTTGGATTGGCACCTGTTAATTACTCACTTGCATTTAAAGAAATTA  768
         | | ||||| ||| || |||| | |||   |||||||| |      |||  ||||| |||
Db   230 ACGCGGTTTATCAAAGCTGGATAACCGGGCATTAATTATTGGCTGATGCTGAAAGAGATTA   289
Qy   769 ATGAGATAACAACAATGCTGGCTTAACAATGATGAGATAAATCATATATCCATTGTTATTTG 828
         | ||| || ||| | | ||| |     |||||| ||  ||| |||| ||   | |||| |||
Db   290 CCGAAGTTGGGCGCCGCAGGGGCGTCGATGTTGGATAAGTAAGTTGTGGGCCTTGCGTTGTGGG 349
Qy   829 GTTTATTTGAAGTTGCTGTGTGTTTTTAAGTTTAAGTTCATAAAAGAAAAGTATATAAAC   888
```

RESULT 12
ACF65385
WP Sequence split into 7 fragments   LOCUS ACF65385 Accession ACF65385
WP   Fragment Name   Begin   End
WP   ACF65385_1   1   110000
WP   ACF65385_1   100001   210000
WP   ACF65385_3   200001   310000
WP   ACF65385_4   300001   410000
WP   ACF65385_5   400001   510000
WP   ACF65385_6   500001   610000
WP   ACF65385_6   600001   618776
ID   ACF65385 standard; DNA; 618776 BP.
XX
AC   ACF65385;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Photorhabdus luminescens nucleotide sequence #38.
XX
KW   Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW   detection; food; gene expression; plant; animal; microorganism; toxin;
KW   antibiotic; biopesticide; virulence factor; disease model; plague;
KW   whooping cough; gene; ds.
XX
OS   Photorhabdus luminescens.
XX
PN   WO200294867-A2.
XX
PD   28-NOV-2002.
XX
PF   07-FEB-2002; 2002WO-IB003040.
XX
PR   07-FEB-2001; 2001FR-00001659.
XX
PA   (INSP) INST PASTEUR.
PA   (CNRS) CNRS CENT NAT RECH SCI.
XX

```
PI   Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI   Buchrieser C;
XX
DR   WPI; 2003-148459/14.
XX
PT   Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT   useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX
PS   Claim 1; SEQ ID NO 38; 1205pp; French.
XX
CC   The invention relates to the isolation of genes and their encoded
CC   proteins from Photorhabdus luminescens. The isolated sequences are
CC   sources of probes and primers for detecting the genome of P. luminescens
CC   and related species; to study polymorphisms; for gene analysis and for
CC   detection/amplification of the genes. Antibodies (Ab) raised against the
CC   polypeptides encoded by the genes are used for detection/identification
CC   of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC   carry a gene-containing vector are used to select compounds that
CC   modulate, regulate, induce or inhibit expression of the genes in plants,
CC   animals or microorganisms other than P. luminescens and are able to alter
CC   response or sensitivity to toxins and antibiotics produced by P.
CC   luminescens. Cells transformed to express the genes are useful for
CC   recombinant production of the proteins, particularly toxins and
CC   antibacterials useful as insecticides, bactericides and fungicides. The
CC   genes, proteins, vectors containing the genes and Ab are also useful
CC   therapeutically (to treat microbial infection by bacteria or fungi that
CC   are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC   biopesticides. Other uses of the genes and the proteins are as virulence
CC   factors and for identifying targets of human diseases for which P.
CC   luminescens is a model (particularly plague and whooping cough). This
CC   sequence represents one of the isolated P. luminescens genes
XX
SQ   Sequence 618776 BP; 181854 A; 139249 C; 125634 G; 172037 T; 0 U; 2 Other;
```

```
Query Match             5.8%;  Score 96.6;  DB 10;  Length 110000;
Best Local Similarity   58.0%;  Pred. No. 5.1e-08;
Matches 171;  Conservative 0;  Mismatches 124;  Indels 0;  Gaps 0;

Qy     37 TAGGTATCGTATTGGCTGCGAGAGGGATATCCAAAGGATTATCGCAAAGGCGATGAAATCA 96
          ||||||  ||| |||| |||||||||| ||||||||| ||||||||
Db  65180 TGGGTGTGGTACTCGCTGCCGGAGGGTATCCTGCTGATTATCACAAAGGGGATATCATTC 65239

Qy     97 GCGGATTGCCTAAAAGTGCGGTCAAAAACGAGAGAAGTTTTCTTAGCGGGTGTCGCAGAAC 156
          ||| ||||||    ||   |||         |||||||||| |||
Db  65240 ATGGCTTGCCCCAACAAGAGAAATGAAAGTAGCAAGGTTTTCCATGCTGGTACAGCAATGA 65299

Qy    157 AAGAGGCAAGCTAGTCACAAACGGCGGTCGTCGTACTTTGTGTGACTGCGTTAGGCGAAA 216
          ||||   |||  ||||  |||||||| || ||||||||  |||| |||| ||||||
Db  65300 AAGACAATGATGTTATTACCGCGGTGGACGCGTTCGTGTGTTACAGCATTAGGAGAAA 65359

Qy    217 GTGTATTTGAAGCACACAAAAAGCGTTAAAATTGGCTGAGCAAATTCAATGGTCTGGGC 276
          |||||   ||| |||||||||||   |||  |||||||   |||||||||||||||||
Db  65360 CCATTATAGATGCACAGAAAAAATGCTTACCAACAAGCAGCAGAGAGATTGAATGGAATAACT 65419

Qy    277 GTTTTTATCGTCGAGACATTGGTTACAGGGCTGTGGAACGAGAACAAGCAAAATA 331
          |||||| |||| ||| ||||||||||||||| ||||||||||||||||||||| |||
Db  65420 GTTTTTACCGCAAGGATATCGGGCAATATCGGGCAATTAACCGATTGAAATAATCATA 65474
```

```
RESULT 13
ACF67367_29
Continuation (30 of 57) of ACF67367 from base 2900001 (Photorhabdus luminescens nucleoti
WP  Sequence split into 57 fragments     LOCUS ACF67367 Accession Acf67367
WP  Fragment Name    Begin      End
WP  ACF67367_00      1          110000
WP  ACF67367_01      100001     210000
WP  ACF67367_02      200001     310000
WP  ACF67367_03      300001     410000
WP  ACF67367_04      400001     510000
WP  ACF67367_05      500001     610000
WP  ACF67367_06      600001     710000
WP  ACF67367_07      700001     810000
WP  ACF67367_08      800001     910000
WP  ACF67367_09      900001     1010000
WP  ACF67367_10      1000001    1110000
WP  ACF67367_11      1100001    1210000
WP  ACF67367_12      1200001    1310000
WP  ACF67367_13      1300001    1410000
WP  ACF67367_14      1400001    1510000
WP  ACF67367_15      1500001    1610000
WP  ACF67367_16      1600001    1710000
WP  ACF67367_17      1700001    1810000
WP  ACF67367_18      1800001    1910000
WP  ACF67367_19      1900001    2010000
WP  ACF67367_20      2000001    2110000
WP  ACF67367_21      2100001    2210000
WP  ACF67367_22      2200001    2310000
WP  ACF67367_23      2300001    2410000
WP  ACF67367_24      2400001    2510000
WP  ACF67367_25      2500001    2610000
WP  ACF67367_26      2600001    2710000
WP  ACF67367_27      2700001    2810000
WP  ACF67367_28      2800001    2910000
WP  ACF67367_29      2900001    3010000
WP  ACF67367_30      3000001    3110000
WP  ACF67367_31      3100001    3210000
WP  ACF67367_32      3200001    3310000
WP  ACF67367_33      3300001    3410000
WP  ACF67367_34      3400001    3510000
WP  ACF67367_35      3500001    3610000
WP  ACF67367_36      3600001    3710000
WP  ACF67367_37      3700001    3810000
WP  ACF67367_38      3800001    3910000
WP  ACF67367_39      3900001    4010000
WP  ACF67367_40      4000001    4110000
WP  ACF67367_41      4100001    4210000
WP  ACF67367_42      4200001    4310000
WP  ACF67367_43      4300001    4410000
WP  ACF67367_44      4400001    4510000
WP  ACF67367_45      4500001    4610000
WP  ACF67367_46      4600001    4710000
WP  ACF67367_47      4700001    4810000
WP  ACF67367_48      4800001    4910000
WP  ACF67367_49      4900001    5010000
WP  ACF67367_50      5000001    5110000
WP  ACF67367_51      5100001    5210000
WP  ACF67367_52      5200001    5310000
WP  ACF67367_53      5300001    5410000
WP  ACF67367_54      5400001    5510000
WP  ACF67367_55      5500001    5610000
WP  ACF67367_56      5600001    5648894
```

```
Query Match             5.8%;  Score 96.6;  DB 10;  Length 110000;
Best Local Similarity   58.0%;  Pred. No. 5.1e-08;
Matches 171;  Conservative 0;  Mismatches 124;  Indels 0;  Gaps 0;

Qy     37 TAGGTATCGTATTGGCTGCGAGAGGGATATCCAAAGGATTATCGCAAAGGCGATGAAATCA 96
          ||||||  ||| |||| |||||||||| ||||||||| ||||||||
Db  70314 TGGGTGTGGTACTCGCTGCCGGAGGGTATCCTGCTGATTATCACAAAGGGGATATCATTC 70373

Qy     97 GCGGATTGCCTAAAAGTGCGGTCAAAAACGAGAGAAGTTTTCTTAGCGGGTGTCGCAGAAC 156
          ||| ||||||    ||   |||         |||||||||| |||
Db  70374 ATGGCTTGCCCCAACAAGAGAAATGAAAGTAGCAAGGTTTTCCATGCTGGTACAGCAATGA 70433

Qy    157 AAGAGGCAAGCTAGTCACAAACGGCGGTCGTCGTACTTTGTGTGACTGCGTTAGGCGAAA 216
          ||||   |||  ||||  |||||||| || ||||||||  |||| |||| ||||||
Db  70434 AAGACAATGATGTTATTACCGCGGTGGACGCGTTCGTGTGTTACAGCATTAGGAGAAA 70493

Qy    217 GTGTATTTGAAGCACACAAAAAGCGTTAAAATTGGCTGAGCAAATTCAATGGTCTGGGC 276
          |||||   ||| |||||||||||   |||  |||||||   |||||||||||||||||
Db  70494 CCATTATAGATGCACAGAAAAAATGCTTACCAACAAGCAGCAGAGAGATTGAATGGAATAACT 70553

Qy    277 GTTTTTATCGTCGAGACATTGGTTACAGGGCTGTGGAACGAGAACAAGCAAAATA 331
          |||||| |||| ||| ||||||||||||||| ||||||||||||||||||||| |||
Db  70554 GTTTTTACCGCAAGGATATCGGGCAATATCGGGCAATTAACCGATTGAAATAATCATA 70608
```

RESULT 14
ACF69918
ID   ACF69918 standard; DNA; 1284 BP.
XX
AC   ACF69918;
XX
DT   20-NOV-2003 (first entry)
XX
DE   Photorhabdus luminescens nucleotide sequence #8385.
XX
KW   Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
KW   detection; food; gene expression; plant; animal; microorganism; toxin;
KW   antibiotic; biopesticide; virulence factor; disease model; plague;
KW   whooping cough; gene; ds.
XX
OS   Photorhabdus luminescens.
XX
PN   WO200294867-A2.
XX
PD   28-NOV-2002.
XX
PF   07-FEB-2002; 2002WO-IB003040.
XX
PR   07-FEB-2001; 2001FR-00001659.
XX
PA   (INSP) INST PASTEUR.
PA   (CNRS) CNRS CENT NAT RECH SCI.
XX
PI   Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
PI   Buchrieser C;
XX
DR   WPI; 2003-148459/14.
XX
PT   Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
PT   useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX
PS   Claim 2; SEQ ID NO 8385; 1205pp; French.
XX
CC   The invention relates to the isolation of genes and their encoded
CC   proteins from Photorhabdus luminescens. The isolated sequences are
CC   sources of probes and primers for detecting the genome of P. luminescens
CC   and related species; to study polymorphisms; for gene analysis and for
CC   detection/amplification of the genes. Antibodies (Ab) raised against the
CC   polypeptides encoded by the genes are used for detection/identification
CC   of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
CC   carry a gene-containing vector are used to select compounds that
CC   modulate, regulate, induce or inhibit expression of the genes in plants,
CC   animals or microorganisms other than P. luminescens and are able to alter
CC   response or sensitivity to toxins and antibiotics produced by P.
CC   luminescens. Cells transformed to express the genes are useful for
CC   recombinant production of the proteins, particularly toxins and
CC   antibacterials useful as insecticides, bactericides and fungicides. The
CC   genes, proteins, vectors containing the genes and Ab are also useful
CC   therapeutically (to treat microbial infection by bacteria or fungi that
CC   are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC   biopesticides. Other uses of the genes and the proteins are as virulence
CC   factors and for identifying targets of human diseases for which P.
CC   luminescens is a model (particularly plague and whooping cough). This
CC   sequence represents one of the isolated P. luminescens genes
XX
SQ   Sequence 1284 BP; 365 A; 263 C; 314 G; 342 T; 0 U; 0 Other;

Query Match   5.8%; Score 95.4; DB 10; Length 1284;
Best Local Similarity   58.7%; Pred. No. 5.8e-08;
Matches 165;   Conservative   0;   Mismatches 116;   Indels   0;   Gaps   0;

Qy   37 TAGGTATCGTATTGGCTGCAGAGGATATCCAAAGGATTATCGCAAAGGCCATGAAATCA 96
          ||||||| ||||||| ||||| |||||| ||||||||||| || | ||||||| |||||
Db   995 TGGGTGTGGTACTCGCTGCCCGAGAGGTATCCTCTGATTATCACAAAGGGGATATCATTC 1054

Qy   97 GCGGATTGCCTAAAAGTGCGGTCAAAAACGAGAAAGTTTCTTAGCGGGTGTCGCGAAC 156
          ||||| ||||| ||| ||||| ||||||||||||||||| || || || ||| |||

Db   1055 ATGGCTTGCCCAACAGAAAATGAAAGTAGCAAGGTTTTCCATGCTGGTACAGCAATGA 1114
          |||| | ||| ||||| |||||| || ||||||||| |||| ||||| | ||||||
Qy   157 AAGAAGGCAAGCTAGTCACAAACGGCGGGTGTCGTGTACTTGTGTGACTGCGAAA 216
          |||| ||||| | ||| |||||| || |||| ||| | |||| ||| |||
Db   1115 AAGACAATGATGTTATTACCCGCCGGTGACGCGTTCTGTGTGTTACAGCATTAGGAGAA 1174

Qy   217 GTGTATTTGAAGCACACAACAAAAAGCGTTAAAAATTGGCTGAGCAAATTCAATGGTCTGGGC 276
          |||| ||||||| ||| ||||| ||| |||||||||| |||||| ||| ||||||||| ||
Db   1175 CCATTATAGATGCACAGAAAAATGCTTACCACAAGCAGAGAGATTGAATGGAAATAACT 1234

Qy   277 GTTTTTATCGTCGAGACATTGTTACAGGGCTGTGAACGA 317
          ||||||| |||| || ||||| |||||| |||| ||||
Db   1235 GTTTTTACCGCAAGGATATCGGATATCGGGCAATTAACCGA 1275

RESULT 15
ADA31630
ID   ADA31630 standard; DNA; 1335 BP.
XX
AC   ADA31630;
XX
DT   20-NOV-2003 (first entry)
XX
DE   DNA encoding Acinetobacter baumannii protein #2917.
XX
KW   ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
KW   vaccine; plant biocontrol agent.
XX
OS   Acinetobacter baumannii.
XX
PN   US652958-B1.
XX
PD   13-MAY-2003.
XX
PF   04-JUN-1999;   99US-00328352.
XX
PR   09-JUN-1998;   98US-0088701P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Breton G, Bush D;
XX
DR   WPI; 2003-576092/54.
DR   P-PSDB; ADA35756.
XX
PT   New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT   for diagnosing a bacterial disease, as components of antibacterial
PT   vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT   plants.
XX
PS   Example; SEQ ID NO 2917; 328pp; English.
XX
CC   The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC   The A. baumannii nucleic acids and polypeptides are useful as reagents
CC   for diagnosing a bacterial disease, as components of antibacterial
CC   vaccines, as targets for antibacterial drugs, to detect the presence of
CC   A. baumannii and other Acinetobacter species in a sample, in screening
CC   compounds for the ability to interfere with the A. baumannii life cycle
CC   or to inhibit A. baumannii infection, and as biocontrol agents for
CC   plants. The present sequence represents DNA encoding an A. baumannii
CC   protein.
XX
SQ   Sequence 1335 BP; 390 A; 240 C; 324 G; 381 T; 0 U; 0 Other;

Query Match   5.8%; Score 95.4; DB 9; Length 1335;
Best Local Similarity   57.6%; Pred. No. 5.9e-08;
Matches 171;   Conservative   0;   Mismatches 126;   Indels   0;   Gaps   0;

Qy   35 ATTAGGTATCGTATTGGCTGCAGAGGATATCGCAAAGGCCATGAAT 94
          |||||| ||||| |||| ||||||||||||| |||||| ||||||
Db   1038 AATCGGTATTGTACTTGCAGCAGAAGGTTACCCAGACAGCGTGCTAAAAGGTGACGTTAT 1097

Qy   95 CAGCGGATTGCCTAAAAGTGCGGTCAAAAACGAGAAAGTTTCTTAGCGGGTGTCGCAGA 154

```
Db   1098 CTCAGGTATCGGTCAATCACCAGGAAGATACTAAAATCTTCCATGCAGGCACTGCTCACTCG 1157
          |   ||  ||   | ||  |      | |        | |      ||        ||
Qy    155 ACAAGAAGGCAAGCTAGTCACAAACGGCGGTCGTCGTTTGTGTGACTTGTGCGTTAGGCGA  214
          ||  |   | || || | |  ||   ||||||||||||||| |||||   |
Db   1158 TGAAGATGGGCATATCGTTACTTCTGGCCGACGTGTACTTTGCGTGACTGCTCTAGGTGA  1217
          ||  |    ||||||||| |||||                                ||
Qy    215 AAGTGTATTTGAAGCACAACAAAAAGCGTTAAAATTGGCTGAGCAAATTCAATGGTCTGG  274
          ||  ||  ||||||||||||   || |  |||  |||        ||        ||
Db   1218 TAGCGTTCTTGAAGCACAAATCAATGCTTTAGAAGTAATGTGGGCAAGTAACCTTCACAGG 1277
           ||  |  |||||||||||||  |  ||  ||    ||| ||||  |||  |||| ||||
Qy    275 GCGTTTTTATCGTCGAGACATTGGTTACAGGGCTGTGGAACGAGAACAAGCAAAATA      331
          ||  |   |||||||||||||||||||| | |   || ||| |||| ||||||||
Db   1278 TATGCAATACCGCAGTGACATTGGTTACCGTGCCATTGCTCGTGAAAAAGCTGAATA      1334
```

Search completed: December 23, 2004, 16:18:53
Job time : 830 secs

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004 Compugen Ltd.

ÖM nucleic - nucleic search, using sw model

Run on:         December 23, 2004, 14:10:34 ; Search time 7035 Seconds
                                    (without alignments)
                             11131.733 Million cell updates/sec

Title:          US-10-698-235-1
Perfect score:  1656
Sequence:       1 aagtaatatcacaatagsgg........ggtgttkggaaraaataacg 1656

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       4526729 seqs, 2364849745 residues

Total number of hits satisfying chosen parameters:    9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      GenEmbl:*
                1: gb_ba:*
                2: gb_htg:*
                3: gb_in:*
                4: gb_om:*
                5: gb_ov:*
                6: gb_pat:*
                7: gb_ph:*
                8: gb_pl:*
                9: gb_pr:*
               10: gb_ro:*
               11: gb_sts:*
               12: gb_sy:*
               13: gb_un:*
               14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1551.6	93.7	1621	1	AY599442	AY599442 Haemophil
2	292.4	17.7	10159	1	U32770	U32770 Haemophilus
3	292.4	17.7	110000	6	AR274513_09	Continuation (10 o
4	288.4	17.4	11122	1	AE006056	AE006056 Pasteurel
5	197.8	11.9	1572	6	AX024062	AX024062 Sequence
6	196.2	11.8	1575	6	A96272	A96272 Sequence 30
7	196.2	11.8	15053	1	AE002550	AE002550 Neisseria
8	196.2	11.8	172325	6	AX044035	AX044035 Sequence
9	196.2	11.8	349980	6	AX044034	AX044034 Sequence
10	185.8	11.2	1575	6	A96276	A96276 Sequence 30
11	185.8	11.2	1575	6	AX024113	AX024113 Sequence
12	134.4	8.1	288108	1	AP005083	AP005083 Vibrio pa
c 13	121.6	7.3	300521	1	AE017153	AE017153 Haemophil
14	120.8	7.3	1275	6	A96274	A96274 Sequence 30
c 15	120.8	7.3	349061	1	NMA222491	AL162753 Neisseria
16	116	7.0	250150	1	AP005342	AP005342 Vibrio vu
c 17	116	7.0	300169	1	AE016801	AE016801 Vibrio vu
18	115.8	7.0	170627	2	AC125567	AC125567 Rattus no
c 19	109	6.6	12116	1	AE004115	AE004115 Vibrio ch

c 20	103	6.2	110000	1	BX950851_02	BX950851_02 Continuation (3 of
c 21	98	5.9	188406	2	BX381547	CR381547 Danio rer
22	97.8	5.9	189598	2	BX957347	BX957347 Danio rer
23	96.6	5.8	110000	1	CR543861_24	CR543861_24 Continuation (25 o
c 24	96.6	5.8	349287	6	BX571860	BX571860 Photorhab
25	96.6	5.8	349980	6	AX770907	AX770907 Sequence
26	95.4	5.8	1335	6	AR320367	AR320367 Sequence
c 27	95.2	5.7	192525	2	BX936428	BX936428 Danio rer
c 28	94.6	5.7	57203	3	AC115581	AC115581 Dictyoste
29	94.6	5.7	181381	2	CR45818	CR45818 Danio rer
c 30	94	5.7	10338	2	AB015493	AB015493 Shewanell
31	94	5.7	349980	6	AX344563	AX344563 Sequence
32	93.8	5.7	183584	9	AC012492	AC012492 Homo sapi
c 33	93.8	5.7	224322	2	CR450711	CR450711 Danio rer
34	92.8	5.6	250029	3	AE014839	AE014839 Plasmodiu
35	92.6	5.6	66993	2	AC138074	AC138074 Homo sapi
36	92.6	5.6	234545	5	BX470214	BX470214 Zebrafish
37	91.8	5.5	157591	2	BX928756	BX928756 Danio rer
38	91.4	5.5	47108	6	AX344506	AX344506 Sequence
c 39	91.2	5.5	67970	3	AL031746	PFMALIP3 Plasmodiu
c 40	91.2	5.5	175282	2	CR352258	CR352258 Danio rer
41	90.6	5.5	903	1	AF064955	AF064955 Coxiella
c 42	90.6	5.5	187161	5	BX08526	BX08526 Zebrafish
c 43	90.4	5.5	131682	3	AL672277	AL672277 Human DNA
44	90.4	5.5	156975	2	CR394534	CR394534 Danio rer
45	90.4	5.5	157141	3	AC016445	AC016445 Drosophil

```
                              ALIGNMENTS

RESULT 1
AY599442
LOCUS       AY599442              1621 bp    DNA     linear   BCT 09-JUN-2004
DEFINITION  Haemophilus influenzae clone L002_HI_0151_004 putative
            phosphoribosylamine-glycine ligase (purD) and hypothetical protein
            genes, partial cds.
ACCESSION   AY599442
VERSION     AY599442.1  GI:48243651
KEYWORDS    .
SOURCE      Haemophilus influenzae
  ORGANISM  Haemophilus influenzae
            Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
            Pasteurellaceae; Haemophilus.
REFERENCE   1  (bases 1 to 1621)
  AUTHORS   Antalis,P., Shen,K., Erdos,G., Gladitz,J., Sayeed,S., Hayes,J.,
            Ahmed,A., Johnson,S., Dice,B., Keefe,R., Dopico,R., Chong,W.,
            Goodwin,J., Singh,M., Janto,B., Post,J.C., Ehrlich,G.D. and Hu,F.Z.
  TITLE     Genomic sequences from a pooled library of 10 clinical isolates of
            Haemophilus influenzae from middle-ear effusions of pediatric
            patients with chronic otitis media
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1621)
  AUTHORS   Antalis,P., Shen,K., Erdos,G., Gladitz,J., Sayeed,S., Hayes,J.,
            Ahmed,A., Johnson,S., Dice,B., Keefe,R., Dopico,R., Chong,W.,
            Goodwin,J., Singh,M., Janto,B., Post,J.C., Ehrlich,G.D. and Hu,F.Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-APR-2004) Center for Genomic Sciences,
            Allegheny-Singer Research Institute, 320 East North Avenue,
            Pittsburgh, PA 15212, USA
FEATURES             Location/Qualifiers
     source          1..1621
                     /organism="Haemophilus influenzae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:727"
                     /clone="L002_HI_0151_004"
     gene            <1..313
                     /gene="purD"
     CDS             <1..313
                     /gene="purD"
                     /codon_start=2
                     /transl_table=11
                     /product="putative phosphoribosylamine-glycine ligase"
```

```
                /protein_id="AAT40791.1"
                /db_xref="GI:48243652"
                /translation="DPRASLGIVLAAEGYPKDYRKGDEISGLPKSAVKNEKVFLAGVA
                EQEGKLVTNGGRVLCVTALGESVFEAQQKALKLAEQIQWSGRFYRRDIGYRAVEREQA
                K"
     misc_feature   90..118
                /gene="purD"
                /note="uptake signal sequence"
     CDS         469..>1621
                /codon_start=1
                /evidence=not_experimental
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="AAT40792.1"
                /db_xref="GI:48243653"
                /translation="MRQFVIVLYSAILLSLEVIYRKLFNISSIERYTESYLSVCLFV
                CLFVFSKYRITRILVGALPAISIVVNNVHYAVYQSWIGPVNYSLAFKEINEITNAGLT
                MIDKFIYPLLFGLFEVAVFLSLSFIKRKVYKLSWIFDFIFYAVMMYVFVRAYTTKSHE
                RFISPNTVYSRLKSNVLSLGYFIGRIVYPEYFLSNIPLYHKSKPMKSGSPKIRNIIL
                IMGESATSHFSAPGYGRKTSPFLDSLKYKSGALVGKTYSGGKLTAISLPMFFNAIPY
                PNGIQIAKGDTNLFNLAKEQGFQTYFYSAQARDDMHMINFLGGAWIDDIRFPDNEGY
                SLRDSMPDNKLJLPAFKNINDNGYHFVVLHHRGSHIPYGALLDEKEKVFG"
     repeat_region   586..613
                /rpt_type=tandem
                /rpt_unit="gttt"
ORIGIN

Query Match    93.7%;   Score 1551.6;  DB 1;   Length 1621;
Best Local Similarity  99.2%;   Pred. No. 4.5e-241;
Matches 1597;  Conservative  4;  Mismatches  4;  Indels  5;  Gaps  4;

Qy      19  GGATCCACGAGCTTCTATTAGTATCGTATTGGCTCGCAGAGGATATCCGAAGGATTATC   78
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db       1  GGATCCACGAGCTTCT-TTAGGTATCGTATTGGCTCGCAGAGGATATCCGAAGGATTATC   59

Qy      79  GCAAAGGCCGATGAAATCAGCCGGATTGCCTAAAAGTGCGGTCAAAAACGAGAAAGTTTCT  138
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      60  GCAAAGGCCGATGAAATCAGCCGGATTGCCTAAAAGTGCGGTCAAAAACGAGAAAGTTTCT  119

Qy     139  TAGCGGGTGTCCGAGAACAAGAGGCAAGCTAGTCACAAAACGGCGGTCGTGTGTACTTGTG  198
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     120  TAGCGGGTGTCCGAGAACAAGAGGCAAGCTAGTCACAAAACGGCGGTCGTGTGTACTTGTG  179

Qy     199  TGACTCGGTTAGGCGGAAAGTGTATTTGAAGCACAACAACAACAAAAAGCGTTAAAATTGGCTGAGC  258
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180  TGACTCGGTTAGGCGGAAAGTGTATTTGAAGCACAACAACAACAAAAAGCGTTAAAATTGGCTGAGC  239

Qy     259  AAATTCAATGGTCTCGGGCGTTTTATCGTCGAGACATTGGTTACAGGGCTGTGGAACGAG  318
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240  AAATTCAATGGTCTCGGGCGTTTTATCGTCGAGACACATTGGTTACAGGGCTGTGGAACGAG  299

Qy     319  AACAAGCAAAATAGTTAGAAATCTTGTTGAAATTAATTAGATAGAAAATATTGTTACAGGG  378
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300  AACAAGCAAAATAGTTAGAAATCTTGTTGAAATTAATTAGATAGAAAATATTGTTACAGGG  359

Qy     379  TAGAATTGTATTTTCCTAGGATTAGGATTTGTTAGGGCAACGTTACGATTGCTCTGA  438
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360  TAGAATTGTATTTTCCTAGGATTAGGATTTGTTAGGGCAACGTTACGATTGCTCTGA  419

Qy     439  CAATAAATTAGAATTATTA-TTTTGTTACTTTATGAGGTTATATCAACTTATGCGACAAT  498
            ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db     420  CAATAAATTAGAATTATTA-TTTTGTTACTTTATGAGGTTATATCAACTTATGCGACAAT  478

Qy     499  TTGTCATCGTAGTATATATTCGCAATTCTTCTTCTTCATTAGAAGTTATTGTCTGTTTGTTTGTTTG  558
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     479  TTGTCATCGTAGTATATATTCGCAATTCTTCTTCTTCATTAGAAGTTATTGTCTGTTTGTTTGTTTG  538

Qy     559  TATTTAATATTTCTAGTATTGAGAGATACACTGAAAGTTATTTGTCTGTTTGTTGTTTG  618
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     539  TATTTAATATTTCTAGTATTGAGAGATACACTGAAAGTTATTTGTCTGTTTGTTGTTTG  598

Qy     619  TTTGTTTGTTTGTTTTTTCAAAATATAGAATTACAAGAATTATTAGTTGGCGCCTTATTG  678
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     599  TTTGTTTGTTTGTTTTTTCAAAATATAGAATTACAAGAATTATTAGTTGGCGCTTATTG   658

Qy     679  CAATAAGTATTGTTGTTAATAATGTACATTATGTGTAATAATGTACATTATGTCAGTACCAATCTTGGATTGGACCTG   738
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     659  CAATAAGTATTGTTGTTAATAATGTACATTATGTGTAATAATGTACATTATGTCAGTACCAATCTTGGATTGGACCTG   718

Qy     739  TTAATTACTCACTTGCATTAAAGAAATTAATGAGATAACACAAATGCTGGCTAACAATGA   798
            |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db     719  TTAATTACTCACTTGCATTAAAGAAATTAATGAGATAACACAAATGCTGGCTAACAATGA   778

Qy     799  TAGATAAATTCATATATCCATTGTTATTGGTTATTGGTTTATTGAAGTTGCTGTGTTTTTTAAGTT   858
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     779  TAGATAAATTCATATATCCATTGTTATTGGTTATTGGTTTATTGAAGTTGCTGTGTTTTTTAAGTT   838

Qy     859  TAAGTTTCATAAAAAGAAAAAGTATATAAAACTTCTTGGATTTTGACTTTATTTTTTATG   918
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     839  TAAGTTTCATAAAAAGAAAAAGTATATAAAACTTCTTGGATTTTGACTTTATTTTTTATG   898

Qy     919  CTGTGATGATGTATGTTTTGTTCGAGCGTATACAACAAAATCCCATGAGCGTTTATTT   978
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     899  CTGTGATGATGTATGTTTTGTTCGAGCGTATACAACAAAATCCCATGAGCGTTTTATTT   958

Qy     979  CACCTAACACTGTTATTCTCGATTAAAATCCAATTATTATCGTTGGGTATTTATTTTATAG  1038
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     959  CACCTAACACTGTTATTCTCGATTAAAATCCAATTATTATCGTTGGGTATTTATTTTATAG  1018

Qy    1039  GACGAATTGTTCCTATGAGATATTTCTTTATCTAAATATTCCTCTTTATCATAAATCTA  1098
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1019  GACGAATTGTTCCTATGAGATATTTCTTTATCTAAATATTCCTCTTTATCATAAATCTA  1078

Qy    1099  AGCCTATGAAATCGGGCTCTCCGGVAAAATTMARGAATAATAATTTAATTAATGGGGAAA  1158
            |||||||||||||||||||||||||  | : |:||||: ||:|||||||
Db    1079  AGCCTATGAAATCGGGCTCTCCGAAA--TTAAGAATATAATTTTTATT-ATGGGGGAAA  1135

Qy    1159  GTGCGACCTCAAGTCATTTTAGTGCTTTAGTGCTTTTGGTTACGGGAGAAAAACATCTCCTTTTTAG  1218
            |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    1136  GTGCGACCTCAAGTCATTTTAGTGCTTTAGTGCTTTTGGTTACGGGAGAAAAACATCTCCTTTTTAG  1195

Qy    1219  ATAGCTTAAAATATAAATCAGGAGCTTCTTGGTAAAAACTTATTCAGGAGGAAAAGCTAA  1278
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1196  ATAGCTTAAAATATAAATCAGGAGCTTCTTGGTAAAAACTTATTCAGGAGGAAAAGCTAA  1255

Qy    1279  CAGCAATTCTTTACCAATGTTTTTACCAATGTTTTAATGCAATYCCTAYCCAAATGGAATACAACAGA  1338
            |||||||||||||||||||||||||| :  | :||||:|||||||
Db    1256  CAGCAATTCTTTACCAATGTTTTTACCAATGTTTTAATGCAATTCCTATCCAAATGGAATACAACAGA  1315

Qy    1339  TAGCTAAAGGAGATACGAATTTATTTATTTATTAGCGAAAGAGCAAGGCTTTCAGACATATT  1398
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1316  TAGCTAAAGGAGATACGAATTTATTTATTTATTAGCGAAAGAGCAAGGCTTTCAGACATATT  1375

Qy    1399  TTTATTCAGCTCAAGCTAGGGATGCGATGATATGATGCATATGATCAATTTTTAGGAGAGCTTGGA  1458
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1376  TTTATTCAGCTCAAGCTAGGGATGCGATGATATGATGCATATGATCAATTTTTAGGAGAGCTTGGA  1435

Qy    1459  TTGATGATATTCGTTTTCCAGATAATGAAGGGTATTCTTAAGAGATTCAATGCCTCGATA  1518
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1436  TTGATGATATTCGTTTTCCAGATAATGAAGGGTATTCTTAAGAGATTCAATGCCTCGATA  1495

Qy    1519  ATAAAATTACTTCCTGCTTTAAAAATATTAATTTAGATAATGGTTATCATTTTGTTTGTTT  1578
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1496  ATAAAATTACTTCCTGCTTTAAAAATATTAATTTAGATAATGGTTATCATTTTGTTTGTTT  1555

Qy    1579  TACATCATAGAGGAGTGCATTAGGGCAATTATTAGAATGAAA  1628
            ||||||||||||||||||||||||||||||||| |||||
Db    1556  TACATCATAGAGGAGTGCATTATCCCTATGGGGCAATTATTAGAATGAAA  1605

RESULT 2
U32770
LOCUS       U32770           10159 bp    DNA     linear   BCT 02-JUN-2004
DEFINITION  Haemophilus influenzae Rd KW20 section 85 of 163 of the complete
            genome.
ACCESSION   U32770 L42023
```

```
VERSION      U32770.1  GI:1573898
KEYWORDS     .
SOURCE       Haemophilus influenzae Rd KW20
  ORGANISM   Haemophilus influenzae Rd KW20
             Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
             Pasteurellaceae; Haemophilus.
REFERENCE    1  (bases 1 to 10159)
  AUTHORS    Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A.,
             Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J., Dougherty,B.A.,
             Merrick,J.M., McKenney,K., Sutton,G.G., FitzHugh,W., Fields,C.A.,
             Gocayne,J.D., Scott,J.D., Shirley,R., Liu,L.I., Glodek,A.,
             Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E.,
             Cotton,M.D., Utterback,T., Hanna,M.C., Nguyen,D.T., Saudek,D.M.,
             Brandon,R.C., Fine,L.D., Fritchman,J.L., Fuhrmann,J.L.,
             Geoghagen,N.S., Gnehm,C.L., McDonald,L.A., Small,K.V., Fraser,C.M.,
             Smith,H.O. and Venter,J.C.
  TITLE      Whole-genome random sequencing and assembly of Haemophilus
             influenzae Rd
  JOURNAL    Science 269 (5223), 496-512 (1995)
  MEDLINE    95350630
   PUBMED    7542800
REFERENCE    2  (bases 1 to 10159)
  AUTHORS    Tatusov,R.L., Mushegian,A.R., Bork,P., Brown,N.P., Hayes,W.S.,
             Borodovsky,M., Rudd,K.E. and Koonin,E.V.
  TITLE      Metabolism and evolution of Haemophilus influenzae deduced from a
             whole-genome comparison with Escherichia coli
  JOURNAL    Curr. Biol. 6 (3), 279-291 (1996)
  MEDLINE    96398784
   PUBMED    8805245
REFERENCE    3  (bases 1 to 10159)
  AUTHORS    White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
  TITLE      Direct Submission
  JOURNAL    Submitted (25-JUL-1995) The Institute for Genomic Research, 9712
             Medical Center Dr, Rockville, MD 20850, USA
REFERENCE    4  (bases 1 to 10159)
  AUTHORS    White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
  TITLE      Direct Submission
  JOURNAL    Submitted (27-SEP-1997) The Institute for Genomic Research, 9712
             Medical Center Dr, Rockville, MD 20850, USA
  REMARK     The H. influenzae sequence has been updated by R. Fleischmann. New
             database matches have been assigned, product names have been
             improved, and a number of frame shifts have been corrected. We
             gratefully acknowledge the work of Tatusov et. al. We have
             incorporated their annotation into the /notes fields of the
             corresponding H. influenzae genes
REFERENCE    5  (bases 1 to 10159)
  AUTHORS    White,O., Clayton,R.A., Kerlavage,A.R., Fleischmann,R.D.,
             Peterson,J., Hickey,E., Dodson,R. and Gwinn,M.
  TITLE      Direct Submission
  JOURNAL    Submitted (28-MAY-1998) The Institute for Genomic Research, 9712
             Medical Center Dr, Rockville, MD 20850, USA
  REMARK     The whole genome was shifted by 588 nucleotides for a new start
             On Sep 30, 1996 this sequence version replaced gi:1221601.
COMMENT      
FEATURES             Location/Qualifiers
     source          1..10159
                     /organism="Haemophilus influenzae Rd KW20"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:71421"
     gene            137..1126
                     /gene="HI0881"
     CDS             137..1126
                     /gene="HI0881"
                     /note="similar to SP:P19641 GB:X68873 PID:388220
                     PID:606125 GB:U00096 percent identity: 64.91; identified
                     by sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="octaprenyl-diphosphate synthase (ispB)"
                     /protein_id="AAC22540.1"
                     /db_xref="GI:1573899"
                     /translation="MKKQDLMSIDEIQKLADPDMQKVNQNILAQLNSDVPLIGQLGFY
                     IVQGGGRHRPLAVLAARSLGFEGSNSITCATFVEFIHTASLLHDDVVDESDMRRGR
                     ATANAEFGNAASVLVGDFIYTRAFQLVAQLESLKILSIMADATNVLAEGEVQQLMNVN
```

DPETSEANYMRVIYSKTARLFEVAGQAAAIVAGGTEAQEKALQDYGRYLGTAFQLVDD
VLDYSANTQALGKNVGDDLAEGKPTLPLLHAMRGHGNAQQAALIREAIEQGGKRRAIDE
VLAIMTEHKSLDYAMNRAKEEAQKAVDAIEILPESEYKQALISLAYLSVDRNY

```
     gene            1177..1914
                     /gene="HI0882"
     CDS             1177..1914
                     /gene="HI0882"
                     /note="hypothetical protein; identified by GeneMark;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="H. influenzae predicted coding region HI0882"
                     /protein_id="AAC22546.1"
                     /db_xref="GI:1573906"
                     /translation="MNTELQPKLEKSAVNFQAKPRKQKIRKDPNAPFIREKLELPDGH
                     NKLLHSCCAPCSGEVMEAILASGIEFTIYFYNPNIHPLKEYLIRKENIRFAKKFGI
                     SRWKDMNQINGCGHRAAEKYDDVIYDYNWRKEGGSQRMEISKRERFYQQEYCGCVY
```

```
     gene            complement(1973..3343)
                     /gene="HI0883"
     CDS             complement(1973..3343)
                     /gene="HI0883"
                     /note="similar to PID:757741 SP:Q45068 PID:1405464
                     GB:AL009126 percent identity: 39.09; identified by
                     sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="amino acid carrier protein, putative"
                     /protein_id="AAC22541.1"
                     /db_xref="GI:1573900"
                     /translation="MTIESILSAIDSFIWGAPLLILLSGTGLYLTLRLGFIQIRYLPR
                     ALGYLFKDKGGKGDVSSFAALCTALAATIGTGNIVGVATAVAQGPGAIFWMLVAL
                     LGMATKYAECLLAVKVRVRDKNGFMAGGPMYIERGLGIRWLAKLFALFGVMVAPFGI
                     GTFPQVNAITHAMQDTFNIPVLVTAIITVLLVGLIILGGVKRIATASSVIVPFMAILY
                     VTTSLVIILLNIEKVPDAILIDSAFDPQAALGGAVGLTVMKAIQSGVARGIFSNES
                     GLGSAPIAAAAQTREPVRQGLISMTGTFLDTIIVCMTGIVLVLTGAWNNPELAGAT
                     VTNYAFAQGLGTSIGATIVTVGLLFFAFTTILGWCYYGERCFVYLVGIRGVKLYRLAY
                     IMLVGLGAFLHLNLIWIIADIVNGLMAFPNLIALIGLRKVIIEETKDYFQRLKINHYD
                     QDEVIK"
```

```
     gene            complement(3741..4451)
                     /gene="HI0884"
     CDS             complement(3741..4451)
                     /gene="HI0884"
                     /note="similar to SP:P03026 GB:L34010 GB:M10044
                     PID:1790863 percent identity: 78.26; identified by
                     sequence similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="aerobic respiration control protein ARCA (arcA)"
                     /protein_id="AAC22542.1"
                     /db_xref="GI:1573901"
                     /translation="MTTPKILVVEDEIVTRNTLKGIFEAEGYDVFEAENGVEMHILA
                     NHNINLVMDINLPGKGLLARELREELSLPILFLTGRDNEVDKILGLEIGADDYLT
                     KFPNPRELTIRARNLLHRAMPHQEKENTFGREFYRFNGWKLDLNSHSLITPEGQEFKL
                     PRSEFRAMLHFCENPGKLQTREELLKKMTGRELKPQDRTVDVTIRRKHFEDHPNTP
                     NIIMTIHGEGYRFCGDIE"
```

```
     gene            4643..6382
                     /gene="HI0885"
     CDS             4643..6382
                     /gene="HI0885"
                     /note="similar to SP:P36655 GB:X77707 GB:236905 PID:536980
                     PID:581055 percent identity: 48.93; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="thiol:disulfide interchange protein (dsbD)"
                     /protein_id="AAC22543.1"
                     /db_xref="GI:1573902"
                     /translation="MKKLFLFFTLIFTTAFAANSGLFDKKQTFLKVDDAFASATLSTD
                     KSQLOAWHDIADGYYLYDGKISAELVGKSNPLSLHTQQAAELHQQDPYFGEVKVFTHSI
                     DGIFRGAFNNADDKVEITYQGCTEGFCYPEFKVLRIIGDLVVSQKQIVEKTVEKNTAL
                     LSEQDRILADGLFHSKWAIFGFFVLGLGLAFTPCVLPMLPLLSAIVIGQQQRPNMRAF
```

```
                                                               Db    9985  ATTGGCTGAGCAAATTCAATGCTCTGGGCGTTTTATCGTCGAGACATTGGTTACAGGGC 10044
SLAFLYVQGMALTYTLLGLAVAAIGLPFQIALQHPYVMIGLSILFVALALSMFGLFTI                       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QLPNSLQNKLNTWSQKQTSGAPGGAFAMGMIAGLVASPCTSAPLSGALLYVAQSGDLF            Qy     308  TGTGGAACGAGAACAAGCAAAATAGTTAGAAATCTTGTTGAATTAATTAGATAAAAAAT  367
TGAVTLYLLALGMGVDLMLITLPKSGEWMNTVKQTFGFVMLALPVFLLSRIIL                            ||||||||||
PEVWEPRLWAGLATVFFIWFALQMSKNGFGYAIKIISFALAMVTVQPLQNWIWQTQTT            Db   10045  TGTGGAACGACGAGTTGCAAAATAATCCATATCATTATGATGAATTATTTCATTTAAGT 10104
TQSAVENMPVSQVKFPKQIKNTEELDRTLAENPHSIAMLDLYADWCVACKEFEKLTFSD                      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
PQVQQQFQNILLQVSMTKNSPENKALMERFNVWGLPTILFFDQQNNEIKGSRVTGFM            Qy     368  ATTGTACAGGGTAGAATTGTATTTCCTAGGATTT  402
DADAFSNWIEKLL"                                                                    ||  ||   |||||  ||
                                                                     Db   10105  GATGAAAATTTAGGAATGAAATTTTAATTAAATTT 10139
gene       6494..6898
           /gene="HI0886"
CDS        6494..6898
           /gene="HI0886"                                          RESULT 3
           /note="hypothetical protein; identified by GeneMark;     AR274513_09
           putative"                                                WPCOMMENT
           /codon_start=1                                           Sequence split into 19 fragments  LOCUS AR274513  Length 110000;
           /transl_table=1                                                Fragment Name       Begin        End
           /product="H. influenzae predicted coding region HI0886"       AR274513_00              1     110000
           /protein_id="AAC22547.1"                                       AR274513_01         100001     210000
           /db_xref="GI:1573903"                                          AR274513_02         200001     310000
           /translation="MNNLEXYRPYFLAFLRIVVAYMFILHGTAKPLEFPISMTCGNGA         AR274513_03         300001     410000
VGDPMLLVAGVIEIVGSILLILGLFTRQAAFILSVEMAYAYFFLHVAGKGNLFFPIAN                   AR274513_04         400001     510000
GGELALLYSLLFLYVFSGAGACALDNKFFKK"                                             AR274513_05         500001     610000
                                                                             AR274513_06         600001     710000
gene       7086..8684                                                        AR274513_07         700001     810000
           /gene="HI0887"                                                    AR274513_08         800001     910000
CDS        7086..8684                                                        AR274513_09         900001    1010000
           /gene="HI0887"                                                    AR274513_10        1000001    1110000
           /note="similar to SP:P15639 GB:J05126 GB:X51950 PID:147420         AR274513_11        1100001    1210000
PID:396345 percent identity: 77.19; identified by sequence                   AR274513_12        1200001    1310000
similarity; putative"                                                        AR274513_13        1300001    1410000
           /codon_start=1                                                    AR274513_14        1400001    1510000
           /transl_table=1                                                   AR274513_15        1500001    1610000
           /product="phosphoribosylaminoimidazolecarboxamide                 AR274513_16        1600001    1710000
formyltransferase (purH)"                                                    AR274513_17        1700001    1810000
           /protein_id="AAC22544.1"                                          AR274513_18        1800001    1830121
           /db_xref="GI:1573904"                                     Continuation (10 of 19) of AR274513 from base 900001 (AR274513 Sequence 1 from patent US
           /translation="MADRPIRQALLSVSDKTGIVEPAQGLVKRGVKLLSTGGTAKLLA         AR274513 Sequence 1 from patent US
QNALPVIEVSDYTGFPEMMDGRVKTLHPKVHGGILGRRGTDRAIMQQHGIEGIDMVVV          AR274513  Accession AR274513
NLYPPAATVAKPDCTLADAVENIDIGGPTMVRSAAKNHKDVAIVVNNHDFNAILAEMD
QHQNSLTFETRFDLAIKAFEHTAQYDSMIANYFGQLVKPYHIAEEEEANAKCCGQFPRT        Query Match       17.7%;   Score 292.4; DB 6;  Length 110000;
LNLNFVRKQAMRYGENSHQNAAFYVDLNVKEASVATANQLGKALSYNNIADTDAALE          Best Local Similarity  85.1%;  Pred. No. 4.9e-38;
CVKEFDDPACVIVKGHANPCGVALGKDILDAYNRAYQTDPTSAFGGIIAFNRELDEKTA        Matches  336;  Conservative   1;  Mismatches   57;  Indels    1;  Gaps      1;
NEIVERQFVEVIIAPKVSAEAQEVMKRKKNVRLLECGEWTSRSRLDFKRVNGGLLVQ
DADLGMVGVDDLKVVSKRQPTEQELKDLLFCWKVAKFVKSNAIVYAKDNQTIGIGAGQ          Qy      8  ATCACAATAGSGGATCCACGAGCTTCTATTAGGTATCGTATTGGCTGCGAGAGGGATATCC   67
MSRVYSAKIAGIKAQDEGLEVAGCVMASDAFFPPRDGIDAAAKVGIQCVIHPGGSMRD                    ||   |:||:|||||||||||||||| ::|||||||:||||||||||||||||||||||||
QEVIDAADEHNMVMVLTGMRHFRH"                                           Db   41919  ATATCACAATGGGATCTGCGAGCTTCT-TTAGGTATCGTATTGGCTGCGAGAGGGATACCC 41977
gene       8780..10069
           /gene="HI0888"                                           Qy     68  AAAGGATTATCGCAAAGGCGATGAAATCAGCGGAGTGCCTAAAAGTGCGGTCAAAAACGA  127
CDS        8780..10069                                                        |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
           /gene="HI0888"                                           Db   41978  AAAAGATTATCGCAAAGGCGATGAAATCAGCGGAGTGCCTAAAAGTGCGGTCAAAAACGA 42037
           /note="similar to SP:P15640 GB:J05126 GB:X51950 PID:147421
PID:396344 percent identity: 75.18; identified by sequence           Qy    128  GAAAGTTTCTTAGGCGGGTGTCGCAGAACAAGAAGCAAGCTAGTCACAAACGGCCGGTCG  187
                                                                              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Query Match       17.7%;   Score 292.4; DB 1;  Length 10159;        Db   42038  GAAAGTTTCTTAGGCGGGTGTCGCAGAACAAGAAGCAAGCTAGTCACAACAAACGGCCGTCG 42097
Best Local Similarity  85.1%;  Pred. No. 7.5e-38;
Matches  336;  Conservative   1;  Mismatches   57;  Indels    1;  Gaps      1;  Qy    188  TGTACTTTGTGTGACTCGGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA  247
                                                                              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      8  ATCACAATAGSGGATCCACGAGCTTCTATTAGGTATCGTATTGGCTGCGAGAGGGATATCC   67  Db   42098  TGTACTTTGTGTGACTCGGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA 42157
           ||   |:||:||||||||||||||| ::|||||||:||||||||||||||||||||||||
Db   9746  ATATCACAATGGGATCTGCGAGCTTCT-TTAGGTATCGTATTGGCTGCGAGAGGGATACCC 9804  Qy    248  ATTGGCTGAGCAAATTCAATGCTCTGGGCGTTTTATCGTCGAGACATTGGTTACAGGGC  307
                                                                              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     68  AAAGGATTATCGCAAAGGCGATGAAATCAGCGGAGTGCCTAAAAGTGCGGTCAAAAACGA  127  Db   42158  ATTGGCTGAGCAAATTCAATGCTCTGGGCGTTTTATCGTCGAGACATTGGTTACAGGGC 42217
           |||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||
Db   9805  AAAAGATTATCGCAAAGGCGATGAAATCAGCGGAGTGCCTAAAAGTGCGGTCAAAAACGA 9864  Qy    308  TGTGGAACGAGAACAAGCAAAATAGTTAGAAATCTTGTTGAATTAATTAGATAAAAAAT  367
                                                                              ||||||||||
Qy    128  GAAAGTTTCTTAGGCGGGTGTCGCAGAACAAGAAGCAAGCTAGTCACAAACGGCCGGTCG  187  Db   42218  TGTGGAACGACGAGTTGCAAAATAATCCATATCATTATGATGAATTATTTCATTTAAGT 42277
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9865  GAAAGTTTCTTAGGCGGGTGTCGCAGAACAAGAAGCAAGCTAGTCACAACAAACGGCCGGTCG 9924  Qy    368  ATTGTACAGGGTAGAATTGTATTTCCTAGGATTT  402
                                                                              ||  ||   |||||  ||
Qy    188  TGTACTTTGTGTGACTCGGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA  247  Db   42278  GATGAAAATTTAGGAATGAAATTTTAATTAAATTT 42312
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9925  TGTACTTTGTGTGACTCGGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA 9984  RESULT 4
                                                                     AE006056
Qy    248  ATTGGCTGAGCAAATTCAATGCTCTGGGCGTTTTATCGTCGAGACATTGGTTACAGGGC  307  LOCUS       AE006056           11122 bp    DNA        linear     BCT 16-JUN-2004
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
DEFINITION  Pasteurella multocida subsp. multocida str. Pm70 section 23 of 204
            of the complete genome.
ACCESSION   AE006056 AE004439
VERSION     AE006056.1 GI:12272442
KEYWORDS    .
SOURCE      Pasteurella multocida subsp. multocida str. Pm70
  ORGANISM  Pasteurella multocida subsp. multocida str. Pm70
            Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
            Pasteurellaceae; Pasteurella.
REFERENCE   1 (bases 1 to 11122)
  AUTHORS   May,B.J., Zhang,Q., Li,L.L., Paustian,M.L., Whittam,T.S. and
            Kapur,V.
  TITLE     Complete genomic sequence of Pasteurella multocida, Pm70
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 98 (6), 3460-3465 (2001)
   MEDLINE   21145866
   PUBMED    11248100
REFERENCE   2 (bases 1 to 11122)
  AUTHORS   Zhang,Q. and Kapur,V.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-2000) Department of Veterinary Pathobiology,
            University of Minnesota, 1971 Commonwealth Ave., St. Paul, MN
            55108, USA
FEATURES             Location/Qualifiers
     source          1..11122
                     /organism="Pasteurella multocida subsp. multocida str.
                     Pm70"
                     /mol_type="genomic DNA"
                     /strain="Pm70"
                     /sub_species="multocida"
                     /db_xref="taxon:272843"
     gene            111..1589
                     /gene="PM0218"
     CDS             111..1589
                     /gene="PM0218"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown"
                     /protein_id="AAK02302.1"
                     /db_xref="GI:12720443"
                     /translation="MKYKDLRDFLTLEQRGELKRIKQEIDPHLEMTEIADRTLRAGG
                     PALLFENPKGFDIPVLCNLFGITPKRVAMGMGQBQVSALRDVGKLLAFLKEPEPPKGFK
                     DFLSSIPQFKQVLNMPTKVLGKADCCQQVVLKDRDVDLYKLPIMQCWKEDVAPLVTWGL
                     TITKGPLKKRQNLGIYRQQLVAKNKLIMRWLSHRGGALDFQEWKETHPGEPPPVSVAL
                     GADPATIIGAVTPVPDTLSEYAPAGLLRGTKTSVVKSVSNDLEVPASAEIVLEGYIDP
                     NETALEGPYGDHTGYYNEQEYFFVFVTVTHITMRKDPIYHSTYTGRPDEPAVLGEALN
                     EVFIPILQKQPFBIVDFYLPPEGCSYRLAVVTIKKQYAGHAKRVMMGVWSFLRQFMYT
                     KFVIVCDDDVNARDWKDVIWAMTRCDPARDLITLVENTPIDYLDFASPVAGLGSKMGI
                     DATNKWPGBTQREWGVPIKKDPDVVKRVDEIWDSLAIPENEP"
     gene            complement(1638..2348)
                     /gene="arcA"
                     /note="synonym: PM0219"
     CDS             complement(1638..2348)
                     /gene="arcA"
                     /codon_start=1
                     /transl_table=11
                     /product="ArcA"
                     /protein_id="AAK02303.1"
                     /db_xref="GI:12720444"
                     /translation="MGTPQILIVEDEAITRNTLKSIFEAEGYEVFERAADGAQMHRILS
                     NKVINLVDINLPGKNGLMLARELRETNTALMFLTGRDNEVDKILGLEIGADDYIT
                     KPFNPRELTIRARNLLQRTMQENSKDSHHPIEQYRFNGWTLDLNSRTLINPEGBEYKL
                     PRSEFRAMLHFCENPGKIQTREBLLKKMTGRELKPQDRTVDVTIRRIRKHFEDHPETP
                     EIIATIHGEGYRFCGELE"
     gene            complement(2736..3002)
                     /gene="rpL31_1"
                     /note="synonym: PM0220"
     CDS             complement(2736..3002)
                     /gene="rpL31_1"
                     /codon_start=1
                     /transl_table=11
                     /product="RpL31"
                     /protein_id="AAK02304.1"
                     /db_xref="GI:12720445"
                     /translation="MKKGIHPENYRTVLFYDSNAKQGFLIRSCARTTTMKWEDGHEY
                     PVFMCDTSSASHPYYTGKTRQIANEGRASDFVNRYGKFGTLKSK"
     gene            3181..4941
                     /gene="dsbD_1"
                     /note="synonym: PM0221"
     CDS             3181..4941
                     /gene="dsbD_1"
                     /codon_start=1
                     /transl_table=11
                     /product="DsbD"
                     /protein_id="AAK02305.1"
                     /db_xref="GI:12720446"
                     /translation="MLSWNDMVLLMKKIFFIFPFLIWTSLSVNAGVFDKKNAFLKABQA
                     FVFTESQQNDRLSLNWDIASGYYLYKKSLEIKGENGQLFIPQLPLAEIHQDDFFGDVE
                     IYRHQLTFSLPFTQLSATKQVEITYQGCTEGFCYPPETRVIALTSLERESGQISSAVL
                     KEBSVLSELPLABQDRLAAKLIESKYAVFWFFLFGLGLAFTPCVLPMLPLLSAIVIGR
                     EQRPSTAKAFGLSVVYVQGMALTYTLLGLIVAAIGLPFQVALQSPYVLLTLSAIFVLL
                     ALSMFGLFTLQLPVSLQTKLTQYSQQQKSGVFGVMGIAGLVASPCTSAGLGAL
                     LVYAQSGDLFTGAITLYLLALGMGLPLPITTVPGNNILPKSGNWMIQVKNAFGFVLLA
                     LPIFLISRVFPEIEMPLWFSLAFAFSLMLLYQFRNNKGVWVFTFILGVMGYLYYSYAM
                     NRHTPAHQTSLSQFERITSYAQLQQVLSKNPNSLAMLDIYADWCVACKFETHTFSD
                     VAVQKAFTDVLLQVDMTKNSEANRELMQKLHVIGLPTLIFFNQQGQEIEGSRITGFM
                     QASPFVDWLQKMKSIQPISQ"
     gene            5114..6715
                     /gene="purH"
                     /note="synonym: PM0222"
     CDS             5114..6715
                     /gene="purH"
                     /codon_start=1
                     /transl_table=11
                     /product="PurH"
                     /protein_id="AAK02306.1"
                     /db_xref="GI:12720447"
                     /translation="MQPNRPIRQALLSVSDKTGIVEFAQALVQRGVKLLSTGGTAKLL
                     ADHGLAVTEVSDYTGFPEMMDGRVKTLHPKVHGGILGRRGTDDEVMSQQGIEGIDMVV
                     VNLYPFAATVAKPNCSLEEAVENIDIGGPTMVRSAAKNHQDVAIVVNNSDFNAILAEM
                     DQHONSLTLETRPDLAIKAFEHTAQYDAMIANYPGQLVKPYFVABEEDAEBAKCGQFPR
                     TLNLNFIRKQTMRYGENGHQKAAPYVEQDVKEASVSTAKQLQGKALSYNNIADTDAAL
                     ECVKSFDEPACVIVKHANPCGVALGADIIAAYNRAYQTDPTSAFGIIAFNRELDAKT
                     AQTIIDRQFVEVIIAPTVABEAKALLKAKKNVRVLECGEWSGTQQRLDVKRVNGGLLV
                     QEADLGMVDLADLKVVSKRQPTEQELKDLLFCWKVAKFVKSNAIVYAKDNQTIGIGAG
                     QMSRVYSAKIAGIKAQDEGLDVAGCVMASDAFFPFRDGIDAAAKVGIQCVIHPGGSMR
                     DQEVIDAADEHNMVMVLTGMRHFRH"
     gene            6754..8340
                     /gene="dcaA"
                     /note="synonym: PM0223"
     CDS             6754..8340
                     /gene="dcaA"
                     /codon_start=1
                     /transl_table=11
                     /product="DcaA"
                     /protein_id="AAK02307.1"
                     /db_xref="GI:12720448"
                     /translation="MKKSLFILLGYSAVLLSEILYRYFKINSLYRISESFLIIFVV
                     LSLFYFSKYRVSRFFIALFFSCGSTLINNVHYEVYQNWINGTNVLLMFKEYWEVTHAGL
                     HMDKLAGGIIWGALDILIFISISRFRQKTHWVADISFILVMGYIFVRSFYTNQBLGI
                     TSNPGYSRIKANFFAFGYFPIGKTLPYDLFNLSNVSVYYRDKPNMTQPPMAKNIVLIMG
                     ESLSASNVGTFGVERQTMPFLDQLVKNAPAETLLKPAYSAGLGTAISLPAFFNAIPRP
                     NGLEQIVSGRTNLFRLAKERGYQTYFYSAQPENQMMIMSGMGKTWVDHLLFPSDIGYK
                     RSEGMHDHALLPLFEQIDLSEGNHFIVLHQRGHAYPAYYLSEBEKAFKENTPLDNYD
                     STLYNTDQPIEKVFKQLKQHHDDWVLIYTSDHCQFVSNQVYNQCTAKEANYLVPIMTY
                     TENTALQQLQRPPLACDRLFHQQLSTFIIKMLGYDMPISDCQHGVINSLLTGDSGYL
                     EVQANQPPAFFIPKNRFKE"
     gene            8343..9632
                     /gene="purD"
                     /note="synonym: PM0224"
     CDS             8343..9632
                     /gene="purD"
                     /codon_start=1
                     /transl_table=11
                     /product="PurD"
                     /protein_id="AAK02308.1"
                     /db_xref="GI:12720449"
```

```
Db            ||||            |      |||.|        :.|  ||
        7355 AACCTAATATGACACAGCCTCC---CATGGCAAAAAATATTGTCCTAATTATGGGAGAAA 7411

Qy      1159 GTTGCGACCTCAGTCATTTTAGTGCTTTTGGTTACGGGAGAAAAACATCTCCTTTTTTAG 1218
             ||| || ||| |||   ||||||||||    ||||| | ||||||||||||||||||||
Db      7412 GTTTGAGTGCGAGCAATGAGACGTTTGGCTATGAAAGGCAAACCATGCCATTTTTGG 7471

Qy      1219 ATAGCTTAAAATATAAATC------AGGAGCTCTTGTTGGTAAAACTTATTCAGGAGGAA 1272
             || || |||  || |      |||||||||||||| ||||||||||||||||
Db      7472 ATCAGTTAGTCAAAAATGCACCGGCAGAAACGCTATTGAAACCAGCCTATTCTGCTGGGT 7531

Qy      1273 AGCTAACAGCAATTCTTACCAATGTTTTAATGCAATYCCTTAYCCAAATGGAATAC 1332
             ||||||| || ||||||||| |:|||   ||||||||: ||||| |||||||
Db      7532 TAGGTACGGCCATTTCATTGCCTGCTCTTTTTAATGCAATTCCTCCGTCCAAATGGTCTAG 7591

Qy      1333 AACAGATAGCTAAAGGAGGATACGAAATTTATTCATTTAATTAGCGAAGAGCAAGGCTTCAGA 1392
             ||||||      || ||||||:|  |||||||||| ||||||||||||||  ||||||
Db      7592 AACAAATTGTTAGTGCGAGAACAAACTTGTTCCGTTAGCGAAGAAGAACGAGGCTATCAGA 7651

Qy      1393 CATATTTTTATTCAGTCAAGCTAGGGATGAATGCATATGCATATGATCAATTTTTAGGAGGAG 1452
             |||||||||||||||||||  ||  ||||||||||||||||:||||||||||||||||
Db      7652 CCTATTTTTATTCAGCGCAGCCTGAAAACCAAATGATGAGTATCAGTGGGAAAA 7711

Qy      1453 CTTGGATTGATGATATTCGTTTCCCAGATAATCGTTTCCAGATAATGAAGATTCAATGC 1512
             ||||||||||||||||||||||||||||||||:|| | |||||||||||||||||||
Db      7712 CCTGGGTTGATCATTATCTTTCCCGAGTGATCATCGGCTATAAACGTTCAGAGGGAATGC 7771

Qy      1513 CTGATAATAAATTACTTCCTGCTTTCCCCCTTTATTTGGAACAAATAGATTGAGTGAAGGAATCATTTG 1572
             || ||||| |   |||||||||| |||||||||:||||||||||| |||| ||:|||:|||||
Db      7772 ATGATCATGCTTACATCAAAGGATCTCACGCACCACTCCTATG 7831

Qy      1573 TTGTTTTACATCATAGAGGAGTCATATTCCCTATG 1608
             || |||||||||||||| || ||||||| ||||||||
Db      7832 TCGTGTTACATCAAAGAGGATCTCACGCACCACCCTATG 7867
```

```
RESULT 5
AX024062
LOCUS        AX024062                1572 bp      DNA       linear     PAT 15-SEP-2000
DEFINITION   Sequence 5 from Patent FR2785293.
ACCESSION    AX024062
VERSION      AX024062.1   GI:10184374
KEYWORDS     .
SOURCE       Neisseria meningitidis
  ORGANISM   Neisseria meningitidis
             Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
             Neisseriaceae; Neisseria.
REFERENCE    1
  AUTHORS    Nassif,X., Tinsley,C., Aujame,L., Perrin,A., Rokbi,B.,
             Bouchardon,A. and Renauld,M.G.
  JOURNAL    Patent: FR 2785293-A 5 05-MAY-2000;
             PASTEUR MERIEUX SERUMS VACC (FR)
FEATURES             Location/Qualifiers
     source          1..1572
                     /organism="Neisseria meningitidis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:487"
     CDS             1..1572
                     /note="unnamed protein product"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="CAC08907.1"
                     /db_xref="GI:10184374"
                     /translation="MKKSLFVLFLYSSLLTASEIAYRFVFGIETLPAAKMAETFALTF
                     MIAALYLFARYKASRLLIAVFFAFSIIANNVHYAVYQSWMTGINYWLMLKEITEVGSA
                     GASMLDKLWLPALWGVLEVMLFCSLAKPHRKTHFSADILFAFLMLIMFVRSPDTKQEH
                     GISPKPTYSRIKANYFSFGYFVGRVLFVQLFDLSRIPAFKQPAPSKIGQGSVQNIVLI
                     MGESESAAHLKLFGYGRETSPFLTRLSQADFKPIVKQSYSAGFMTAVSLPSFNAIPH
                     ANGLEQISGGDTNMFRLAKEQGYETYFSAQAENEMAILNLIGKKWIDHLIQPTQLGY
                     GNGDNMPDEKLLPLPFDKINLQQGRHFIVLHQRGSHAPYSALLQPQDKVFGELIVDKYD
                     NTIHKTDMIQTVFEQLQKQPDGNWLFAYTSDHGQYRQDIYNQGTVGNLITDAG
                     YSSNKAVQQAANOAFAPCEIAFHQQLSTFLIHTLGYDMPVSGCREGSVTGNLITGDAG
                     SLNIRDGKAEYVYPQ"
```

```
Qy     ORIGIN

Query Match       17.4%;    Score 288.4;    DB 1;    Length 11122;
Best Local Similarity  56.1%;    Pred. No. 3.3e-37;
Matches  626;    Conservative  4;    Mismatches 465;    Indels 21;    Gaps 4;

Qy       499 TTGTCATCGTAGTATTATATTCTGCAATTCTTCTTCATTAGAAGTTATTATAGAAAAT 558
             ||| ||| |   || ||||||||||||   | |||||||    |||| |||||||||
Db      6767 TTTTTATTTTATTGGGCTATCCGCCGTTTTGCTATTGAGCGAAATACTTTATCGTATT 6826

Qy       559 TATTTAATAATTTCTAGTATTGAGAGATACACTGAAAGTTATTTGTCGTTTGTTTTGTTG 618
             |||||||||   || |||||||||||||  || || ||||||||||||||||||||||||
Db      6827 TTTCAAGATTAATAAGCTTATACCGTTATACCGTATGAAAGTTTCTGATTATCTTGTTGTCC 6886

Qy       619 TTTGTTTGTTGTTTTCAAAATATAGAATTACAAGAATATTAGTTGGCGCTTTATTG 678
             |||||| ||| |||||  |||||| | | |||  ||||||||| ||||||||||||
Db      6887 TTTCATTATTTTATTTTTCTAAATATAGGGTATCACGCTTTTTTATCGCTTTTATTCTTT 6946

Qy       679 CAATAAGTATTGTTGTTAATAATGTACATTATGTCAGTATACCAATCTTGGATTGGACCTG 738
             ||   || |||||| | | ||||| |||| |||||||||||||||  || ||||||||||
Db      6947 CTTGTAGTACGTTGATTAATAACGTGCATTATGAGGTTTATCAAAATGAATAATGGCA 7006

Qy       739 TTAATTACTACCTGCATTTAAAGAAATTAATGAGATAAGAAATTAATGAAATAAGGCTGGCCTAACAATGA 798
             ||  | |  |||||  ||  || |||| ||| |  |  || | ||| ||||| ||||||   | ||
Db      7007 CGAATTATCTACTAAGTTAAAGAAATATTGGGAAGTCACCCGCGGTTTACATATGT 7066

Qy       799 TAGATAAATTCATATATCCCATTGTTATTGGTTTATTGTTTGGGGGCGCTAGA-------TATCCTGATTT 858
             | ||||| |  |||| | || |  | |||| ||  |||||| ||||||||      ||||||||||
Db      7067 TAGACAAACTTGGATCGGGCGTTATTGGTTGTGCTTTATCT 7117

Qy       859 TAAGTTTCATAAAGAAAAGTATATAAATCCAATTATTTTGACTTTATTTTTATG 918
             ||||| |||  ||||| || || | |||  |  |||| |||   || ||||||||
Db      7118 TTATTTCCATATCCGGCGTTTATCTTTGTGCGTTATCGTTTGATATTTCGTTTTATCT 7177

Qy       919 CTGTGATGATGATGTTTTTGTTGTCGAGCGTGGCATTATTTTGCGCTATACCACCATGAGGCGTTTTATTT 978
             || || |   ||| || ||  |||||||| ||| |  || ||    |||| |||||| |  |||| |||
Db      7178 TGGTGATGGGCTATATCTTTGTGCCGTTTCTTCCATTAATTGGGAGTTAGGGGATCA 7234

Qy       979 CACCTAACACTGTTATTCTGATTAAAATCCGATTAAAATCCAATTATTTATCGTGGTTATTTTATTATAG 1038
             ||||| | | |||||||  ||  | || ||| || ||  |||  | | || | |||| | ||  ||||
Db      7235 CCTCTAATCCTGGATATTCGAGAATTAAAGCCAACTTTTTCGCCTTTTGCCCCTTGTTACTTTATTG 7294

Qy      1039 GACGAATTGTTCCTTATGAGAGATATTTCTTATTCCTCTTTATCATAAATCTA 1098
             | |||  ||  |||||| | | |||||| ||||||||  |||   |||| |||||
Db      7295 GTAAAACACTGCCTTATGATCCTCTTAATTTGAGTAATGTGTCTGTATTATCGAGATA 7354

Qy      1099 AGCCTATGAAATCGGGCTCTCCGVAAAATTMARGAATATAATTTAATTAATGAGGGGAAA 1158
```

us-10-698-235-1.rge

ORIGIN

```
Qy    1606 ATGGGGCATTATT 1618
           | ||||| || ||
Db    1115 ACAGGGCATTGTT 1127
```

RESULT 6
A96272

LOCUS   A96272   1575 bp   DNA   linear   PAT 07-SEP-2000
DEFINITION   Sequence 305 from Patent WO9924578.
ACCESSION   A96272
VERSION   A96272.1 GI:6780028
KEYWORDS   .
SOURCE   unidentified
  ORGANISM   unidentified
             unclassified.
REFERENCE   1
  AUTHORS   Pizza,M., Scarlato,V., Rappuoli,R., Grandi,G. and Masignani,V.
  TITLE   Neisserial antigens
  JOURNAL   Patent: WO 9924578-A 305 20-MAY-1999;
             PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
             CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
FEATURES   Location/Qualifiers
     source   1..1575
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

Query Match   11.8%;   Score 196.2;   DB 6;   Length 1575;
Best Local Similarity 51.5%;   Pred. No. 3.6e-22;
Matches   532;   Conservative   2;   Mismatches 481;   Indels   18;   Gaps   3;

Qy   589 CTGAAAGTTATTTGTCTGTTTGTTTGTTTGTTTGTTTTCAAAATATAGAA 648
          ||||| |||||||||||||||||||||| || ||| ||||||||
Db   110 CGGAAACGTTTGCGCTGACATTTGTGATTGCTGCGCCTGTATCTGTTTGCGCGTTATAAGG 169

Qy   649 TTACAAGAATATTAGTTGGCGCTTATTTGCAATAAGTATTGTTGTTAATAAATGTACATT 708
          ||||| ||| | | || || || ||||| ||| ||| ||||||||
Db   170 TGACGCGTTTGTTGATTGCGGTGTTTTTCGGTTCAGCATTATTGCCAACAATGTGCATT 229

Qy   709 ATGCAGTATACCAATCTTGGATTGACCTGTTAATTACTCACTTGCATTTAAAGAAATTA 768
          |||| ||| |||||| ||| |||||| || |||||| || | |||||| |||||
Db   230 ACGCCGTTTATCAAAGTTGGATTGGCGAAAGCGGCGGGCATCAATTATTGGCTGAAAGAGGTTA 289

Qy   769 ATGAGATAACAAAATGCTGGCTTAACAATGATAGATAAATTCATATATCCATTGTTATTTG 828
          || | || ||||||| | ||| ||||||| ||| ||||||| ||||||| |||||
Db   290 CCGAAGTCGGCAGTGCAGTGCGCGGCGTCGATGTTGGATAAGTTGTGGCGTGTCGTGTGGG 349

Qy   829 GTTTATTTGAAGTTGCTGCTGTGTTTTCATAAAAAGAAAAGTATATAAAC 888
          ||||| ||||||||| || || |||| |||| | ||||||||| ||||||
Db   350 GCGTGTTGGAAGTCATGTGTTTGTTTGTCGAGCCTGCCAAGTTCCACCGTAAGACGCGCATT--- 405

Qy   889 TTTCTTGGATTTTTGACTTATTTTTATGCTGTGATGATGTTTTTGTTGAGCGT 948
          ||||| |||||||| | ||| || || |||| ||| || || |||||||
Db   406 -------TTTTCTGCCGATATACTGTTTGCCTTCCTAATGCGTGTCCTTCCTGATGTTTTCGTGTCGGTT 457

Qy   949 ATACAACAAAATCCCATGAGGCGTTTTATTTCCACCTAACACTGTTTATTTCTCGATTAAAAT 1008
          |||| |||||| | |||||||||||| ||||| ||| || ||| || |||| ||
Db   458 CGTTCGACACGAAAGACGCGTATTTCGCCCAAACCGACATACAGCCGCATCAAAG 517

Qy   1009 CCAATTATTATCGTTGGGTTATTTTATAGGACGAATTGTTCCTATGAGATATTTCTT 1068
           ||||||||||||||| || |||| ||| |||| | | | || | ||||||||||||
Db   518 CCAATTATTTCAGCTTCGGTTATTTTGTCGGATTATTTGTCGGTATCAGTTGTTTGATT 577

Qy   1069 TATCTAATATTCCTCTTTATCATAAATCGGGCTCCCGGTAAGCCTTTAAGCAGCCTTT 1128
           || || | | | || || ||| || ||| | | ||| | ||||||||
Db   578 TAAGCAGGATTCCCGGCTTTAAGCAGCCTTTCAGATGTCATTTTAGTGTGCTTTTG 634

Qy   1129 MARGAATATAATTTTAATTAATCGGTTACATCATAGAGGGAGTCATATTCCCT 1188
           | ||| | | | | | | | | | | |||||| |||||| || ||| || |
Db   635 TTCAAAATATCGTCCTGATTATGGGCCATTTTTAGTGCTTTTG 694

Qy   1189 GTTACCGGAGAAAAACATCTCCTTTTTTAGATAG---CTTAAAATATAAATCAGGAGCTC 1245
           ||| | | | | | | |||| | ||||| || || ||| |||||
Db   695 GCTACGGACGCGAAACTTCGCCGGTTTTACCCGCTCGTGTCAACCGATTTTAAGCCGA 754

Qy   1246 TTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTCTTACCAATGTTTTTTTA 1305
           |||||||||| ||| |||| ||| || || || ||| | | | ||||| ||||
Db   755 TTGTGAAACAAAGTTATTCCGCAGGCTTTATGACTGCAGTCTGCCCAGTTTTTTCA 814

Qy   1306 ATGCAATYCCTTAYCCAAAATGAATACAACACAGTAGCTAAAGAGAGATACGAATTTATTA 1365
           |||||| || | || ||| ||| |||| || ||||||| | ||||| ||| |||
Db   815 ATGCGGATACCGCACGCCAACGCGGCGGATACTAATAATGTGTTCC 874

Qy   1366 ATTTAGCGGAAAAGAGCAAGGCTTTCAGAACATATTTTTATTCAGCTCAAGTAGGGATGATA 1425
           || | ||| |||||||||| ||| ||| || |||| ||||| || || ||||||||| |
Db   875 GCCTCGCCCAAAGAGCAGCAGCGGGCTATGTATTTTTACAGCGCACAGGCGGAAAACGAGA 934

Qy   1426 TGCATATGATCAATTTTGGATTGATGATATTCGTTTTCCAGATAATG 1485
           ||||||| | | ||| ||| |||| ||||| || ||| |||| ||
Db   935 TGGGGATTTTTGAACTTAATCGGTAAGAAGAAATGGATAGAACCATCTGATTCAGCCAGTG 994

Qy   1486 AAGGGTATTCTTTAAGAGATTCAATGCCATCTGATATAATAAAATAACTTCCTGCTTTAAAAATA 1545
           |||| ||| || | | | | || ||| | || ||||| ||
Db   995 TTGGCTACGGCAACGGCAATATGCCCGGATAGAAGCTGCGTGTTCGCCGTGTTCGACAAAA 1054

Qy   1546 TTAATTAGATAATGGTTATCATTTGTGTTTTACATCATAGAGGGGAGTCATATTCCCT 1605
           |||| | ||| | | |||| | ||| | ||| |||| | | ||| | |||
Db   1055 TCAATTTGCAGCAGGGACAGGCAGCCATTTTATCGTTGTTCACCAACGTGGTTGCACGCCCCAT 1114
```

```
Qy   1189  GTTACGGGAGAAAAACATCTCCTTTTTTAGATAG---CTTAAAATAAATCAGGAGCTC  1245
                ||||| ||| ||| | |||||||     |||||   ||| ||||  || |||||||
Db    695  GCTACGGACGACGAAACTTCGCCGTTTTAACCCGGCGTCGCAAGCCGATTTAAGCCGA  754

Qy   1246  TTGTTGGTAAAACTTATTCAGGAGGAGAAAGCTAACACAGCAATTTCTTTACCAATGTTTTTA  1305
                ||| |||   ||| ||||  ||| ||||| | ||||| || ||  ||| |||  ||||
Db    755  TTCTGAAACAAAGTTATTCCGCAGGCGTTTATGACTGCAGTGTCCCTGCCCAGTTTTTTCA  814

Qy   1306  ATGCAATYCCTAYCCAAATGAATACAACAGTAGCTAAAGGAGATACGAATTTATTTA  1365
                |||| ||| ||  || || || |||||||||||| |||  || |||| |||| || ||
Db    815  ATGCGATACCGCACGCCAACGGCTTGGAACAAATCAGCGCGGCGATACCAATATGTTCC  874

Qy   1366  ATTTAGCGAAAGACAAGCAAGGCTTTCAGACATATTTTATTCAGCTACGGGATGATA  1425
                |||| ||||| |||| ||| ||| |||||| |||| ||| |||||| |||||||||
Db    875  GCCTCGCCAAAGACGAGCAGGGCCATGAAACGTATTTTTACAGCCGCCGAAAAACGAGA  934

Qy   1426  TGCATATGATCAATTTTTTAGGAGGAGCTTGATTGATGATATTCGTTTTCCAGATAATG  1485
                ||| ||||| |||||||   |||| |||   || ||| || ||   ||| ||||| ||
Db    935  TGGCGATTTTGAACTTAATCGGTAAGAAAATGGATAGACCATCTGATTCAGCCGACGCAAC  994

Qy   1486  AAGGGTATTCTTTAAGAGATTCAATGCCTGATAATAAATTACTTCCTGCTTTAAAAATA  1545
                |  |||| ||| ||||| || ||| |||||| ||||| ||| |||  ||||  ||||
Db    995  TTGGCTACGGCAACGGCGACAATATGCCCGATGAGAAGTCGCTGCCGTTGTTCGACAAAA  1054

Qy   1546  TTAATTTAGATAATGGTTATCATTTTGTTGTTTTACATCATAGAGGGAGTCATATTCCCT  1605
                ||||| ||| || || ||| ||||| || ||| |||  || ||||| |||| ||||||
Db   1055  TCAATTGCAGCAGGGCAAGCATTTTATCGTGTTGCACCAACGCGGTTCGCACGCGCCCAT  1114

Qy   1606  ATGGGGCATTATT  1618
                |  |||||| ||
Db   1115  ACGGGCGCATTGTT  1127


RESULT 7
AE002550
LOCUS       AE002550              15053 bp    DNA     linear   BCT 25-MAY-2000
DEFINITION  Neisseria meningitidis serogroup B strain MC58 section 192 of 206
            of the complete genome.
ACCESSION   AE002550 AE002098
VERSION     AE002550.2  GI:7413476
KEYWORDS    .
SOURCE      Neisseria meningitidis MC58
  ORGANISM  Neisseria meningitidis MC58
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
REFERENCE   1  (bases 1 to 15053)
  AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
  TITLE     Complete genome sequence of Neisseria meningitidis serogroup B
            strain MC58
  JOURNAL   Science 287 (5459), 1809-1815 (2000)
   MEDLINE   20175755
   PUBMED    10710307
REFERENCE   2  (bases 1 to 15053)
  AUTHORS   Tettelin,H., Saunders,N.J., Heidelberg,J., Jeffries,A.C.,
            Nelson,K.E., Eisen,J.A., Ketchum,K.A., Hood,D.W., Peden,J.F.,
            Dodson,R.J., Nelson,W.C., Gwinn,M.L., DeBoy,R., Peterson,J.D.,
            Hickey,E.K., Haft,D.H., Salzberg,S.L., White,O., Fleischmann,R.D.,
            Dougherty,B.A., Mason,T., Ciecko,A., Parksey,D.S., Blair,E.,
            Cittone,H., Clark,E.B., Cotton,M.D., Utterback,T.R., Khouri,H.,
            Qin,H., Vamathevan,J., Gill,J., Scarlato,V., Masignani,V.,
            Pizza,M., Grandi,G., Sun,L., Smith,H.O., Fraser,C.M., Moxon,E.R.,
            Rappuoli,R. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
```

```
COMMENT     On Apr 4, 2000 this sequence version replaced gi:7227263.
FEATURES             Location/Qualifiers
     source          1..15053
                     /organism="Neisseria meningitidis MC58"
                     /mol_type="genomic DNA"
                     /strain="MC58"
                     /db_xref="taxon:122586"
                     /note="serogroup: B"
     gene            complement(74..982)
                     /gene="NMB2000"
     CDS             complement(74..982)
                     /gene="NMB2000"
                     /note="conserved hypothetical protein; identified by
                     Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAF42327.1"
                     /db_xref="GI:7227264"
                     /translation="MNQTAINRADVRTRPIFDDMPVRGLHVRLENVWQHIVKQKNYPA
                     AIRRALGELLAAGVLLSGNLKNEGTLIVQVQGRGRLKMLVAEAASDRTVRATARWDET
                     AEIADDESLGDLLGEGGVVFVLTLQPKDGEPWQGVVPLEGGGIAQMLVNYMKRSEQLDT
                     HIVLSASDEAAGLLVQRLPEEVLDEEAWEHVSTLARTLTAEELAGLDAQHVLYRLFH
                     ETPPRVFEPETFEFSCTCSRGKVSDMLLMLGGEEVGGVVVEQGSIEVDCDFCHSKYVF
                     DETDVNALFGEDVGVAKGLPRHTVQ"
     gene            1227..1982
                     /gene="NMB2001"
     CDS             1227..1982
                     /gene="NMB2001"
                     /note="conserved hypothetical protein; identified by
                     Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAF42328.1"
                     /db_xref="GI:7227265"
                     /translation="MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQPAED
                     EQPVLPINRAPARRAGNADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLN
                     EQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAM
                     GINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTG
                     KNIEITSLSHKYWSGKYAFARRVKNDPSRFLN"
     gene            2001..2219
                     /gene="NMB2002"
     CDS             2001..2219
                     /gene="NMB2002"
                     /note="hypothetical protein; identified by Glimmer2;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="AAF42329.1"
                     /db_xref="GI:7227266"
                     /translation="MSMPEMPKWYDDDQIVSCTEKVKVMSENMAELYQTAQDAFEDA
                     LLMCCGEQRLRDYLLALIEGLENPYRKV"
     gene            2364..2711
                     /gene="NMB2003"
     CDS             2364..2711
                     /gene="NMB2003"
                     /note="conserved hypothetical protein; identified by
                     Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAF42330.1"
                     /db_xref="GI:7227267"
                     /translation="MNIRALLILGLCLATGETAVFLAGIKLPGSIVGMGVLFALLQA
                     GWVKTSWLQQLTDALMSNLTLFLVPPCVAVISYIDLIADDWFSILVSASASTLCVLLV
                     TGKVHRWIRGIRR"
     gene            2708..3400
                     /gene="NMB2004"
     CDS             2708..3400
                     /gene="NMB2004"
                     /note="conserved hypothetical protein; identified by
```

NSYTHLLNDPDWLLQYLFAQQNNIAVANNQKLNRNEKVTNSKTKLDILQEAWETLLPH
RKLHITADDIQVSAVDNEELYSASNMSDGERALFYILGQVLSVDDGSVLIFDEPELHI
HKSIISNLWDKIEBELRPDCSFLIITHDIEPAATRVAKKVVIRNYPTPAWDISEVPES
NFDEETITMILGSRKPILFVEGNNNSLDIATYRYCYPDWTIIPKGACKDVIQSVSSLK
KLSNEMPLLNLKCSGIVDLDSRDERBIEQLNNLGIYILPVSEIENLFSLTDVAKEILK
LNQYSDEELLNKLNGFKSELIKYIDNELKDDKLDEFVVKQVRRKIDNYLKNIDLSSKI
TSTDMKKSLLNEISTLTEQKIETWISEIKNEIQRCIEQQPLDKLLTIYDNKGLLAKSA
CVLKGMRNKHEFESWIMRTLKGRNKPFIDAIRQKLPILD"
9568. .9750
/gene="NMB2009"
/note="This region contains a gene with one or more
premature stops or frameshifts, and is not the result of a
sequencing artifact; similar to SP:Q09530 PID:733602
percent identity: 61.70; identified by sequence
similarity; putative"
9962. .11539

Query Match   11.8%;   Score 196.2; DB 1;   Length 15053;
Best Local Similarity   51.5%;   Pred. No. 2.4e-22;
Matches   532; Conservative   2; Mismatches   481; Indels   18; Gaps   3;

```
Qy     589 CTGAAAGTATTTGTCTGTTTGTTGTTTGTTGTTTGTTGTTTTCAAAATATAGAA  648
           ||||  | || ||||| ||| |  || ||| |||||||  || |  || |||||
Db   10074 CGGAAACGTTTGCGCTGACATTTGATTGCTGCGCGTATCTGTTGCGCGTTATAAGG 10133

Qy     649 TTACAAGAATATTAGTTGGGCCTTTATTGCAATAAGTATTGTGTTAATAATGTACATT  708
           |    | |  | || || |   || | |||| || | || || |    ||  ||| ||
Db   10134 TGACCGCGTTTGTTGTTTGGCGTGTTTTTGCGTTCAGCATTATTGCCAACAATGTGCATT 10193

Qy     709 ATGCAGTATACCAAACTTGGATTGGACCTGTTAATTACTCACTTGCATTAAAGAAATTA  768
           ||||| ||  | | ||||  ||| || |  |  | ||| |  || ||| | ||  || |
Db   10194 ACGCGGTTTATCAAAGCTGGATGACGGGCCATCAATTATTGGCTGATGCTGAAAGAGGTTA 10253

Qy     769 ATGAGATAACAAAATGCTGGCTTAACAATGCTGGATAAATTCATATATCCATTGTTATTTG  828
           | || | |||| | || ||||| |||| ||   || |||||  |   |    || |||||
Db   10254 CCGAAGTCGGCAGCCGCGGGGTGCGTCGATGTTGGATAAGTTGGTGGCTGCCTGTTGTTGTGGG 10313

Qy     829 GTTTATTTGAAGTGCTGTGTTTTAAGTTTCATAAAAGTATATAAAC            888
           ||||| ||||||| || |||| |  |  | || ||||     |||
Db   10314 GCGTGTTGGAAGTCATGTGTTTTGCAGCCTGCCAAGTTCCGCGTAAGACGCAT----       10369

Qy     889 TTTCTTGGATTTTGACTTTATTTTTATGCTGTGATGATGTCTTCCTTCGTGTTTTGTTCGAGCGT  948
           ||||| |     |  ||  ||| ||||| |||| | ||||  || | ||| |||| ||||||||
Db   10370 ------TTTTCTGCCGATATACTGTTTGCCTTCGTATGGATGATTTTCGTGCGTT 10421

Qy     949 ATACAACAAATCCCATGAGCGTTTTATTTCACCTAACACTGTTTATTCTCGATTAAAAT 1008
           | || ||| |  | | |   ||  | |||||    || || |||  | || | || |||
Db   10422 CGTTCGACACGAACAAGAGCACGGTATTTCGCCCAAACCGCCATCAAAG 10481

Qy    1009 CCAATTATTATCGTTGGGTTATTTTATAGGACGAATTGTTCCTTATGAGATATTTTCTT 1068
           |||| |||| | ||  |||||| |||     ||| | |||||||  ||| |||||| ||
Db   10482 CCAATTATTTCAGCTTCGGTTATTTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATT 10541

Qy    1069 TATCTAATATTCCTCTTATCATAAATCTAAGCCTATGAAATCGGCGCTCCGVAAAATT 1128
           | |  |  || | || || |  || |  |  ||  | |  |  |||   |   | |||
Db   10542 TAAGCAGGATTCCCGCCTTTAAGCAGCCTGCTCCCAGCAAGCAAAATCGGG---CAGGGCAGTG 10598

Qy    1129 MARGAATATAATTTTAATTAATGGGGAAAGTGCGACCTCAAGTCATTTAGTGCTTTTG 1188
           ||| |   |   || ||     | || || |||     | ||  |||||| |||||||
Db   10599 TTCAAAATATCGTCCTGATTATGGGCGAAAGCGAAAGCGCGGCCATCAATTGAAGCTGTTTG 10658

Qy    1189 GTTACGGGGAGAAAAACATCTCCTTTTTTAGATAG---CTTAAAATATAAATCAGGAGCTC 1245
           ||| | |||   |  || ||| |||||    ||||   |||| |  |||| |||| ||| |
Db   10659 GCTACGGACGCGAAACTTCGCCGTTTTTAACCCGGCTGTCGCAAGCCGATTTTAAGCCGA 10718

Qy    1246 TTGTTGGTAAACTATTCAGGAGGAAAGCTAACAGCAATTTCTTTACCAATGTTTTTTA 1305
           || || |  |  | |||||||   | || |  || |||   | ||||  |  |||||
Db   10719 TTGTGAAACAAAGTTATTCCCGCAGCGCTTTATGACTGCAGTCGTGTCCCGCCCAGTTTTTCA 10778

Qy    1306 ATGCAATYCCTTAYCCAAATGGAATACAACAGATAGTCAACAAGATACGAATTTATTA 1365
           ||||  ||| |   |  ||  |   || |   | ||    |  | ||| |   ||||||
Db   10779 ATGCCGATACCCGCACCCAACCGGCTTGGAACAAATCAGCGGCGGCCGGATACCAACAATATGTTCC 10838
```

gene   Glimmer2; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAF42331.1"
                     /db_xref="GI:7227268"
                     /translation="MNEILROPSVLLFLTLAVYALAIIVRTRTGNIFCNPVLVSTIVL
                     IAYLKILGIDYAVYHNAAQFIDFWLKPAVVVLAVPLYQNRRKIFNQWLPVIVSQLAGS
                     VTGIVTGMYFAKKWIGAEREVVLSLASKSVTNPIAIEITRSIGGIPAITAATVIIAGLV
                     GQIAGYKMLKNTVVMPSSVGMSLGTASHAMGIAASLERSRRMAAYACGLGLTPNGVLTA
                     LIAPLLIPVLGF"
gene   3467. .4687
            /gene="NMB2005"
CDS   3467. .4687
            /gene="NMB2005"
            /note="similar to GB:M65216 SP:P38434 GB:585363 PID:150243
            PID:246444 percent identity: 98.77; identified by sequence
            similarity; putative"
            /codon_start=1
            /transl_table=11
            /product="glutamate N-acetyltransferase/amino-acid
            acetyltransferase"
            /protein_id="AAF42332.1"
            /db_xref="GI:7227269"
            /translation="MAVNLTEKTAEQLPDIDGIALYTAQAGVKKPGHTDLTLIAVAAG
            STVGAVFTNRFCAAPVHIAKSHLFDEDGVRALVINTGNANAGTGAQRIDALAVCAA
            AARQIGCKPNQVLPFSTGVILEPLPADKIIAALPKMQPAFWNEAARAIMTTDFVPKAA
            SREGKVGDKHTVRATGIAKGSGMIHPNMATMLGFIATDAKVSQPVLQLMTQEIADETF
            NTITVDGDTSTNDSFVIIATGKNSQSEIDNIADPRYAQLKEILCSLAIELAQAIVRDG
            EGATKKFITVRVENAKTRDEARQAAYAVARSPLVKTAFFASDPNLGRILLAAIGYAGVAD
            LDTDLVEMYLDDILVAEHGGRAASYTEAQGQAVMSKABITVRIKLHRGQAAATVYTCD
            LSHGYVSINADYRS"
gene   4747. .6156
            /gene="NMB2006"
CDS   4747. .6156
            /gene="NMB2006"
            /note="similar.to PID:1652876 percent identity: 52.04;
            identified by sequence similarity; putative"
            /codon_start=1
            /transl_table=11
            /product="chloride channel protein-related protein"
            /protein_id="AAF42333.1"
            /db_xref="GI:7227270"
            /translation="MTSTFPRRLARKIRQTRRLSRKSIAFLFLLAGSALVALTALFFA
            HLADFALELNAKLVQQYPFAWVALPLGLPLIAWLTRKPAPFTAGSGIPQVIASLSLP
            YGAQKTRLIRLGQTLLKIPLTFLGMLFGASIGREGPSVQVGAAVMGAWGAWCKKHGLA
            FKGMQENDLMAAGAAGGLAAAFNAPLAGVFIFAIEELGRGIMLRWERQILLGVLASGFI
            QVAIQGNNPYFSGFNGGVLEHIFLWVALSGLVCGAAGGLFGRLLYRGAAAFAPRKIRG
            FIRNRPLLAALMGLLLALLGTFYQGKTYGTGVHEAAQALHGIYEAPFGLAAAKWLAT
            VFSYWAGVPGGIFTPSLTIGAVLGEHIAAIADISQGANIIVLICMAAFLAGATQSPIT
            SAVVVMEMTGGOSLLFWMLIACIFASQVSRQFSPRPFYHASGMRFRQRVLQETAAQTG
            NAPARPQTANSKTGMPSEN"
gene   6346. .7737
            /gene="NMB2007"
            /note="This region contains a match to at least one other
            gene that is not full length,and is not the result of a
            sequencing artifact; similar to GP:1742299 percent
            identity: 79.36; identified by sequence similarity;
            putative"
gene   7881. .9524
            /gene="NMB2008"
CDS   7881. .9524
            /gene="NMB2008"
            /note="similar to PID:862629 percent identity: 53.03;
            identified by sequence similarity; putative"
            /codon_start=1
            /transl_table=11
            /product="ABC transporter, ATP-binding protein-related
            protein"
            /protein_id="AAF42334.1"
            /db_xref="GI:7227271"
            /translation="MDEIQIPKKVELQTKLENEKIVLSKGSTTIIVGANGTGKTRLAV
            YIEEQLKEKAHRISAHRALKLNPNVNKIPEKSAKTYLSYGQNWDGIDVSNRKNYRWDN

Qy   829 GTTTATTGAAGTTGCTGTGTTTTAAGTTTAAGTTTCATAAAAGAAAAGTATATAAAC   888
Db 23426 GGCGTGTTGGAAGTCATCGTGTTTGTTTGCAGCCTTGCCAAGTTCCGCCGTAAGACGCCAT--- 23481

Qy   889 TTTCTTGGATTTTGACTTATTTTTATGCTGTGATGATGTATGTTTTTGTTCGAGCGT   948
Db 23482 ------TTTTCGCCGATATACTGTTTGCCTTCCTAATGCTGATGATTTTCGTGCGTT 23533

Qy   949 ATACAACAAAATCCCATGAGCGTTTATTTATTCACCTAACACTGTTATTCTCGATTAAAAT 1008
Db 23534 CGTTCGACACGAAACAAGAGACACGGTATTCGCCCAAACGACATACAGCCGCCATCAAAG 23593

Qy 1009 CCAATTATTATCGTTGGTTATTTTGTTGTTTTACATCATAGGACGAATTGTTCCTTATGAGATATTTCTT 1068
Db 23594 CCAATTATTCAGCTTCGGTTATTTTGTTGCGACGCGTGTTGCGCCGATCAGTTGTTTGATT 23653

Qy 1069 TATCTAATATTCCTCTTTTATCATAAATCTAAGCCTATGAAATCGGGCTCTCCGVAAAATT 1128
Db 23654 TAAGCAGGATTCCGCCCTTTAAGCAGCCTGCTCAAGCAAAATCGGG---CAGGGCAGTG 23710

Qy 1129 MARGAATATAATTTAATTAATGGGGAAAGTGCGAAAGCGAAAGCGCGGCGCATTTGAGAGCTGTTTG 1188
Db 23711 TTCAAAATATGTCCTGATTATGGGCAAAGCGAAAGCGCGATAG---CTTAAAATATAAATCAGGAGCTC 23770

Qy 1189 GTTACGGGAGAAAAACATCTCCTTTTTAGATAG---CTTAAACCCGGCTGTGCCAAGCCGATTTTAAGCCGA 1245
Db 23771 GCTACGACGCGAAACTTCGCCGTTTTTAACCCGGCTGTGCCAAGCCGATTTTAAGCCGA 23830

Qy 1246 TTGTTGTAAACTTATTCAGGAGGAAAGCTAACAGCAATTTCTTTACCAATGTTTTTTA 1305
Db 23831 TTGTGAAACAAGTTATTCCGCAGGCTTTATGCTGCAGTGCTGCAGTGTCCCCTGCCCCAGTTTTTTCA 23890

Qy 1306 ATGCAATYCCTTAYCCAAATGAAATACAACAGATAGCTAAAGGAGATACGAATTATTATTA 1365
Db 23891 ATGCGATAACCGCACGGCAACGGCTTGGAACAAATCAGCGGCGGCGATACCAATATGTTCC 23950

Qy 1366 ATTTAGCGAAAGAGCAAGGCTTTCAGACATATTTTTATTCAGCTCAAGCTAGGGATGATA 1425
Db 23951 GCCTCGCCAAAGAGCAGCAGGGCTATGAAACGTATTTTTACAGGCGCGGAAAACGAGA 24010

Qy 1426 TGCATATGATCAATTTTTTAGGAGGAGCTTGGATTGATGATATTCGTTTCCAGATAATG 1485
Db 24011 TGGGCGATTTTGAACTTAATCGTAAGAAAATGGATAGACCATCTGATTCAGCCGACGCAAC 24070

Qy 1486 AAGGGTATTCTTTAAGAGATTCAATGCCTGATAATAATAATTACTTCCTGCTTTAAAAATA 1545
Db 24071 TTGGCTACGGCAACGGCACAATATGCCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAA 24130

Qy 1546 TTAATTTAGATAATGGTTATCATTTTGTGTTTTACATCATAGAGGGAGTCATATTCCCT 1605
Db 24131 TCAATTTGCAGCAGGCAAGCATTTATCGTTGCGTTGCCACCAACGCGCGTTCGCACGCGCCCCAT 24190

Qy 1606 ATGGGGCATTATT 1618
Db 24191 ACGGGCGCATTGTT 24203

RESULT 9
AX044034
LOCUS   AX044034   349980 bp   DNA   linear   PAT 24-NOV-2000
DEFINITION   Sequence 113 from Patent WO0066791.
ACCESSION   AX044034
VERSION   AX044034.1 GI:11342918
KEYWORDS   .
SOURCE   Neisseria meningitidis
  ORGANISM   Neisseria meningitidis
             Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
             Neisseriaceae; Neisseria.
REFERENCE   1
  AUTHORS   Pizza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
             Masignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,
             Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.
  TITLE   Neisseria genomic sequences and methods of their use

---

Qy 1366 ATTTAGCGAAAGAGCAAGGCTTTCAGACATATTTTTATTCAGCTCAAGCTAGGGATGATA 1425
Db 10839 GCCTCGCCAAAGAGCAGCAGGGCTATGAGGAGGAGCTTGGATTGATGATATTCGTTTCCAGATAATG 10898

Qy 1426 TGCATATGATCAATTTTTTAGGAGGAGCTTGGATTGATGATATTCGTTTCCAGATAATG 1485
Db 10899 TGGCGATTTGAACTTAATCGGTAAGAAATGGATAGACCATCTGATTCAGCCGACGCAAC 10958

Qy 1486 AAGGGTATTCTTTAAGAGATTCAATGCCTGATAATAATAATTACTTCCTGCTTTAAAAATA 1545
Db 10959 TTGGCTACGGCAACGGCGACAATATGCCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAA 11018

Qy 1546 TTAATTTAGATAATGGTTATCATTTTGTGTTTTACATCATAGAGGGAGTCATATTCCCT 1605
Db 11019 TCAATTTGCAGCAGGCAAGCATTTATCGTTGCGTTGCCACCAACGCGCGTTCGCACGCGCCCCAT 11078

Qy 1606 ATGGGGCATTATT 1618
Db 11079 ACGGGCGCATTGTT 11091

RESULT 8
AX044035
LOCUS   AX044035   172325 bp   DNA   linear   PAT 24-NOV-2000
DEFINITION   Sequence 114 from Patent WO0066791.
ACCESSION   AX044035
VERSION   AX044035.1 GI:11342919
KEYWORDS   .
SOURCE   Neisseria meningitidis
  ORGANISM   Neisseria meningitidis
             Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
             Neisseriaceae; Neisseria.
REFERENCE   1
  AUTHORS   Pizza,M., Hickey,E., Peterson,J., Tettelin,H., Venter,J.C.,
             Masignani,V., Galeotti,C., Mora,M., Ratti,G., Scarselli,M.,
             Scarlato,V., Rappuoli,R., Frazer,C.M. and Grandi,G.
  TITLE   Neisseria genomic sequences and methods of their use
  JOURNAL   Patent: WO 0066791-A 114 09-NOV-2000;
             CHIRON CORPORATION (US) ; THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES   Location/Qualifiers
     source   1..172325
                     /organism="Neisseria meningitidis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:487"
                     /note="sequence too long, cut in 8 pieces.-seq 1: 1 to
                     349980 349980 bases-seq 108: 300001 to 649980 349980
                     bases-seq 109: 600001 to 949980 349980 bases-seq 110:
                     900001 to 1249980 349980 bases-seq 111: 1200001 to 1549980
                     349980 bases-seq 112: 1500001 to 1849980 349980 bases-seq
                     113: 1800001 to 2149980 349980 bases-seq 114: 2100001 to
                     2272325 172325 bases"
ORIGIN

Query Match   11.8%; Score 196.2; DB 6; Length 172325;
Best Local Similarity 51.5%; Pred. No. 1.6e-22;
Matches 532; Conservative 2; Mismatches 481; Indels 18; Gaps 3;

Qy   589 CTGAAAGTTATTTGTCTGTGTTTGTTTGTTTTGTTTTTGTTGTTTTCAAAATATAGAA 648
Db 23186 CGGAAAACGTTTGCGCTGCACATTTGTGATTGCTGCGCTGTATCGTGTTTGCGGTTATAAGG 23245

Qy   649 TTACAAGAATATTAGTTGGCGCCTTATTTGCAATAAGTATTAGTGTTAATAAGTGTACATT 708
Db 23246 TGACGCGCGTTTTGTTGTGTGTTTTGCGGTGTTTTTTGCCTTCAGCAATTATTGCCAACAATGTGCATT 23305

Qy   709 ATGCAGTATACCAAATCTTGGAGTTGGACCTGTTAATTACTCACTTGCATTTAAAGAAATTA 768
Db 23306 ACGCCGGTTTATCAAAGCTGGATGACGGGCATCAATTATTGGCTGATGATGATGCTGAAAGGAGGTTA 23365

Qy   769 ATGAGAGATAACAACAAATGCTGGCTTAACAATGATAGAATAAATTCATAATATCCATTGTTATTTG 828
Db 23366 CCGAAGTCGGCAGCGCGGCGGTGCGTCGATGTGGGATAAGTTGTGGCTGCCTGTTGTTGTGGG 23425

JOURNAL   Patent: WO 0066791-A 113 09-NOV-2000;
             CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES   Location/Qualifiers
     source   1..349980
                  /organism="Neisseria meningitidis"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:487"
                  /note="sequence too long, cut in 8 pieces.~seq 1: 1 to
                  349980 349980 bases~seq 108: 300001 to 649980 349980
                  bases~seq 109: 600001 to 949980 349980 bases~seq 110:
                  900001 to 1249980 349980 bases~seq 111: 1200001 to 1549980
                  349980 bases~seq 112: 1500001 to 1849980 349980 bases~seq
                  113: 1800001 to 2149980 349980 bases~seq 114: 2100001 to
                  2272325 172325 bases"

Query Match   11.8%;   Score 196.2; DB 6; Length 349980;
Best Local Similarity   51.5%;   Pred. No. 1.4e-22;
Matches 532;   Conservative   2;   Mismatches 481;   Indels 18;   Gaps 3;

Qy   589 CTGAAAGTTATTTGTCTGTTTGTTTGTTTGTTTGTTTGTTTTTTCAAAATATAGAA 648
            |||||| ||| || ||||| | |||| | |||||||| |||| ||||||||||
Db   323186 CGGAAACGTTTGCCTGACATTTGATGATGCTGCCGTGTCTCGTTTGCGCGTTATAAGG 323245

Qy   649 TTACAAGAATATTAGTTGGCGCTTATTTGCAATAAGTATTGTTGTTAATAATGTACATT 708
            |||| | || |||| || ||||||||||| |||||| |||||| ||| ||||||
Db   323246 TGACGCGTTTGTTGATTGCGGTGTTTTTTTCGTTCAGCATTATGCCAACAATGTGCATT 323305

Qy   709 ATGACAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTTGCATTTAAAGAAATTA 768
            ||||| |||| ||||| |||| ||| || ||||| |||| |||| ||||||||||||
Db   323306 ACGCGGTTTATCAAAGCTGGATGACGGGCATCAATTATTGGCTGATGCTGAAGAGGTTA 323365

Qy   769 ATGAGATAACAAATGCTGCTTAACAATGATAGAAAATCATATATCCATTGTTATTTG 828
            || ||||||| || | || | ||| | | ||| | | |||| |||||| |||
Db   323366 CCGAAGTCGCAGCGCGGGGTGCCGATGTGGTAAGTTGTGGCTGTTGTCGTGTGTGGG 323425

Qy   829 GTTTATTTGAAGTTGCTGTGTTTTTAAGTTTCATAAAAGAAAAGTATATAAAC 888
            || ||||| |||| || ||||||| |||| ||| |||| ||||| |||||
Db   323426 GCGTGTTGGAAGTCATGTGTTTTTGCCAGCCTTGCCAAGTTCCGCGTAAGACCGCAT--- 323481

Qy   889 TTTCTTGGATTTTGACTTTATTTTTATGCTGTGTGATGATGATGTATGTTTTGTTCGGCGTT 948
            |||||||||| ||| || |||| |||||| |||| |||| |||||| |||| ||||
Db   323482 -------TTTCTGCCGATATACTGTTTGCCTCCTAATGCTGAAACTGTTATTCTGCGATTAAAAT 323533

Qy   949 ATACAACAAAATCCCATGAGCGCGTTTTATTTCACCTAACACTGTTTATTCTGCGATTAAAAT 1008
            |||| | | | | |||| ||||||||||||||| | ||||| ||| ||||||| ||
Db   323534 CGTTCGACACGAAACAAGAGCACACGGTATTTCGCGCCAAAACACAGCAGCCGCATCAAAG 323593

Qy   1009 CCAATTATTTATCGTGGGTTATTTATAGGACGAATTGTTCCTTATGGACGAATATTTCTT 1068
            | |||||||| | || |||||||| | ||| |||| ||||| ||| |||||||||||
Db   323594 CCAATTATTTCAGCTGCAGACTTTGTCTGGACGCGTGTCGCCGGTATCAGTTGTTTGATT 323653

Qy   1069 TATCTAATATTCCTCTTATCATAAATCAAGCTAACAGCAATTTCTTACCAATGTTTTTA 1128
            ||| | | |||| | | ||| || |||||||| | |||| | ||| || || |
Db   323654 TAAGCAGGATTCCCGCCTTAAGCAGCCTCCCAAGCAAAATCGGG---CAGGGCAGTG 323710

Qy   1129 MARGAATATAATTTTAATTAAATGGGAGAAGTGCGACCTCAAGTCATTTTAGTGCTTTTG 1188
            ||| |||| || ||||||| ||||||||| | | ||| ||||||||||| ||| |||
Db   323711 TTCAAAATAGTCGTCCTGATTATGGCGAAAGGCGGGCCATTTGAAGCTGTTTG 323770

Qy   1189 GTTACGGAGAAAAACATCTCCTTTTTTTAGATAG---CTTAAAATATAAATCAGGAGCTC 1245
            || |||||| |||| ||| || ||||| ||| | |||||| ||| || || ||||
Db   323771 GCTACGGACGAAACTTCGGCGTTATGCTTCGCAGCGCTGACTGCAGTGTCCCCTGCCCAGTTTTTCA 323830

Qy   1246 TTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTTCTTACCAATGTTTTTA 1305
            ||| | ||||| | |||||| | ||| || ||||| ||||||||||||||| ||||
Db   323831 TTGTGAAACAAAGTTATTCCGGACGCTTTATGACTGCAGTGTCCCAGTTTTTTCA 323890

Qy   1306 ATGCAATYCCTTAYCCAAATGAATAACAACADAGATAGCTAAAGGAATACGAACTTATTA 1365
            |||||| | | | ||||||| || | ||||| | | |||||| | | || ||||
Db   323891 ATCGATACCGCACGCCAACACGCTTGGACAAATCAGCGGCGGCGGATACCAACATGTTCC 323950

Qy   1366 ATTTAGCGAAAGAGCAAGGCTTTCAGACATATTTTATTCAGCATGGGATGATGATA 1425
            | |||

Db   Db (top region, Qy/Db pairs):
DB   323951 GCCTCGCCAAAGAGCAGCAGGGCTATGGAAACGTATTTTTACAGCGCGCAGGCGGAAAACGAGA 324010

Qy   1426 TGCATATGATCAATTTTTGGATGGAGAGCTGGATTGATGATATTCGTTTCCAGATAATG 1485
            ||| |||

Db   324011 TGGCGATTTTGAACTAATCGGTAAGAAATGGATAGAACCATCTGATTCAGCGACGCAAC 324070

Qy   1486 AAGGGTATTCTTAAGAGATTCAATGCCTGATAATAAATTACTTCCTGCTTTAAAAATA 1545
            ||

Db   324071 TTGGCTACCGGCAACGGCGACAATATGCCCGATGAGAAGCTGCTGCCGTTGTTCGACAAAA 324130

Qy   1546 TTAATTAGATAAGTAATGTTATCATTTGTGTTTTACATCATAGAGGAGTCATATTCCCT 1605
            |||||

Db   324131 TCAATTTGCAGCAGGCAAGCATTTTATCGTGTCGCACCAACGCGCGTTCGCACGCGCCCAT 324190

Qy   1606 ATGGGGCATTATT 1618
            ||||| ||||||
Db   324191 ACGGGCGCATTGTT 324203

RESULT 10
A96276
LOCUS   A96276   1575 bp   DNA   linear   PAT 07-SEP-2000
DEFINITION Sequence 309 from Patent WO9924578.
ACCESSION   A96276
VERSION   A96276.1   GI:6780030
KEYWORDS   .
SOURCE   unidentified
  ORGANISM unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Pizza,M., Scarlato,V., Rappuoli,R., Grandi,G. and Masignani,V.
  TITLE   Neisserial antigens
  JOURNAL   Patent: WO 9924578-A 309 20-MAY-1999;
            PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
            CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
FEATURES   Location/Qualifiers
     source   1..1575
                  /organism="unidentified"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:32644"
ORIGIN

Query Match   11.2%;   Score 185.8; DB 6; Length 1575;
Best Local Similarity   50.7%;   Pred. No. 1.7e-20;
Matches 524;   Conservative   4;   Mismatches 487;   Indels 18;   Gaps 3;

Qy   589 CTGAAAGTTATTTGTCTGTTTGTTTGTTTGTTTGTTTGTTTTTTCAAAATATAGAA 648
            |||||| ||| || ||||| | |||| | |||||||| |||| ||||||||||
Db   110 CGGAAACGTTTGCCTGACATTTGATGATGCTGCCGTGTCTCGTTTGCGCGTTATAAGG 169

Qy   649 TTACAAGAATATTAGTTGGCGCTTATTTGCAATAAGTATTGTTGTTAATAATGTACATT 708
            |||| | || |||| || ||||||||||| |||||| |||||| ||| ||||||
Db   170 CTTCGCGGCTGCGTGTTTTTGCGGTGTTTTTTTCGTTCAGCATTATGCCAACAATGTGCATT 229

Qy   709 ATGACAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTTGCATTTAAAGAAATTA 768
            ||||| |||| ||||| |||| ||| || ||||| |||| |||| ||||||||||||
Db   230 ACGCGGTTTATCAAAGCTGGATGACGGGTATTAACTGATGATGAAGCTGAAGAGGTTA 289

Qy   769 ATGAGATAACAAATGCTGGCTTAACAATGATAGAAAATCATATATCATTGTTATTTTG 828
            || ||||||| || | ||| ||| | | ||| | | |||| |||||| |||
Db   290 CCGAAGTCGGCGGCGGCGGCGCGTCGATGTGGTGGATAAGTTGTGGCCTGGCCTTTGTGTGGG 349

Qy   829 GTTTATTTGAAGTTGCTGTGTTTTTAAGTTTCATAAAAGAAAAGTATATAAAC 888
            || ||||| |||| || ||||||| |||| ||| |||| ||||| |||||
Db   350 GCGTGTGGCGGAAGTCATGTGTGTTTGCCAGCCTTGCCAAGTTCCGCGTAAGACGCAT---- 405

Qy   889 TTTCTTGGATTTTGACTTTATTTTTATGTCGTGTGATGATGTATGTTTTTGTTCGGCGTT 948
            |||||||||| ||| || |||| |||||| |||| |||| |||||| |||| ||||
Db   406 --------TTTTCTGCCGATATACTGTTTGCCTCCTAATGCTGAAACTGTTATTCTGCGATTAAAAT 457

Qy   949 ATACAACAAAATCCCATGAGCGGCGTTTTATTTCACCTAACACTGTTTATTCTGCGATTAAAAT 1008
            |||| | | | | |||| ||||||||||||||| | ||||| ||| ||||||| ||
Db   949 ATACAACAAAATCCCATGAGCGGCGTTTCAGACATATTTTATTCTGCGATTAAAAT

```
/db_xref="GI:10184425"
/translation="MKKSLFVLFLYSSLLTASEIAYRFVFGIETLPAAKMAETFALTF
MIAALYLFARYKASRLLIAVFFAFSMIANNVHAVYQSWMTGINYWLMLKEVTEVGSA
GASMLDKLWLPALMGVAEVMLFCSLAKFRRKTHFSADILFAFLMLMIFVRSFDTKQEH
GISPKPTYSRIKANYFSFGYFVGRVLPYQLFDLSKIPVFKQPAPSKIGQGSIQNIVLI
MGESESAAHLKLFGYGRETSPFLTRLSQADFKPIVKQSYSAGFMTAVSLPSFNVIPH
ANGLEQISGDTNMFRLAKEQGYETYFYSAQAENQMAILNLIGKKWIDHLIQPTQLGY
GNGDNMPDEKLLPLFDKINLQQGRHFIVLHQRGSHAPYGALLQPDKVFGEADIVDKY
DNTIHKTDQMIQTVFEQLQKQPDGNWLFAYTSDHGQYVRQDIYNOGTVQPDSYIVPLV
LYSPDKAVQQAANQAFAPCEIAFHQQLSTFLIHTLGYDMPVSGCREGSVTGNLITGDA
GSLNIRNGKAEYVYPQ"
```

ORIGIN

```
Query Match        11.2%;   Score 185.8;  DB 6;  Length 1575;
Best Local Similarity  50.7%;  Pred. No. 1.7e-20;
Matches  524;  Conservative  4;  Mismatches  487;  Indels  18;  Gaps  3;
```

```
Qy   589  CTGAAAGTTATTGTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT--TTTCAAAATATAGAA  648
          |||||   ||| || ||||| |||| ||| || | |||  | | |  |||||||||
Db   110  CGGAAACGTTTGCGTGACATTTATGATTGCTGCGCGTATCGTTTGCGGCTTATAAGG  169
```

```
Qy   649  TTACAAGAATATTAGTTGGCGCTTTATTTGCAATAAGTATTGTTGTTAATAATGTACATT  708
          |||   |||||||| ||| || | |||||| |||||
Db   170  CTTCGCGGCTGCTGATGGTGTTTTCGGTGTTTCGGCGTTCAGCATGATTGCCAACAATGTGCATT  229
```

```
Qy   709  ATGCAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTTGATTTAAAGAAATTA  768
          |||   ||||||||||||      ||||||| |
Db   230  ACGCGGTTATCAAAGCTGGATGACGGGTATTAACTAATTGCTGATGCTGAAAGAGGTTA  289
```

```
Qy   769  ATGAGATAACAAATGCTGGCTTAACAAATGATGATAAATTCATATATCCAATTGTTATTTG  828
          ||| ||| ||  |||  |  ||||| ||||||     ||||| || | |
Db   290  CCGAAGTCGGCAGCGCGGGCGCGTCGAATGTTGGATAAGTTGTGGCTGCTTTGTCGTGGG  349
```

```
Qy   829  GTTTATTTGAAGTGCTGCTTTGTGTTTTCATAAAAGAAAAGTATATAAAC  888
          || | |||||   ||||| ||||  ||| | || ||| | |||| ||
Db   350  GCGTGGCGGAAGTCATGTTGGTTTGCCAAGTCGCCAAGTCCGCGCCGTAAGACGCAT---  405
```

```
Qy   889  TTTCTTGGATTTTTATTTTTTATGCTGTGATGAATGTGTTTTGTTGGAGCGT  948
          |||||||||
Db   406  ------TTTTCTGCCGATATACTGTTTGCCTTCCACTGTCGATGCTGATGATTTTCGTGCGTT  457
```

```
Qy   949  ATACACACAAAATCCCATGAGCGTTTTATTCACCTAACACACTGTTTATTCTGCGATTAAAAT  1008
          |||||||||||| |||||||| |||| |
Db   458  CGTTCGACACGCAAACAAGAGCACACGGTATTTTCGCCCAAACCGACATACAGCCCGCATCAAAG  517
```

```
Qy   1009  CCAATTATTATTCGTTGGGTTATTTATCATAAATCTAAGCCTATGAAATCGGGCTCTCCGVAAAATT  1068
           ||||||||||||||||||
Db   518  CCAATTATTTCAGCTTCGGTTATTTATTTGTCGGGCGCGTGTTGTTGCCGTATCAGTTCAGTTGTTT  577
```

```
Qy   1069  TATCTAATATTCCTCTTTATCATAAATCTAAGCCTATGAAATCTAAGCCTCCGVAAAATT  1128
           ||||| ||||||||||
Db   578   TAAGCAAGATCCCTGTGTTCAAACAGCCTGCTCCAAGCAAAATCGGGCAGGCAGTATTC  637
```

```
Qy   1129  MARGAATATATAATTTTAATTAATGGGGAAAGTGCGACCTCAAGTCATTTTAGTGTCTTTTG  1188
           |  | |   |||||
Db   638   AAAATATCGTCCTG---ATTATGGGCGAAAGCGAAAGCGCGGCGCCATTTGAAATTGTTTG  694
```

```
Qy   1189  GTTACGGGAGAAAAACATCTCCTTTTTTAGATAG---CTTAAAATATAAATCAGGAGCTC  1245
           |||||||
Db   695   GTTACGGGCGGCGAAACTTCGCCGTTTTAACCCGGCCGTGTCGCCAAGCCGATTTTAAGCCGA  754
```

```
Qy   1246  TTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTTCTTTACCAATGTTTTTTA  1305
           ||||||||||
Db   755   TTGTGAAACAAAGTTATTCCGCAGGCTTATGACGGCAGTATCCCTGCCCAGTTCTTTA  814
```

```
Qy   1306  ATGCAATYCCTTAYCCAAATGAATACAACAGATAGCTAAGGATACGAATTTATTTA  1365
           |||| ||  | |||| | | || |  |||  ||||| ||| ||| ||
Db   815   ACGTCATACCCGCACGCAACCGCTTGGAACAAATCAGCGGCGGCGATACCAATATGTTCC  874
```

```
Qy   1366  ATTTAGCGAAAGAGCAAGGCTTTCAGACATATTTTATTCAGCTCAAGCTAGGGATGATA  1425
           || |||||  || | |  | |||   ||||||| | ||||||
Db   875   GCCTCGCCAAAGAGCAGCAGGAGGCGTATTTTTTACAGTGCCCAGGCTGAAAACCAAA  934
```

```
Qy    1426  TGCATATGATCAATTTTTAGGAGGAGCTTGGATTGATGATAATTCGTTTCCAGATAATG  1485
                  ||||||  |||||||||| ||||  ||| ||||||||||||||||||||||||||||
Db     935  TGGCAAATTTGAACTTAATCGGTAAGAAATGGATAGACCATCTGATTCAGCCGACGCAAC  994

Qy    1486  AAGGGTATTCTTTAAGAGATTCAATGCCTGATATAAATTACTTCCTGCTTTTAAAAATA  1545
                  || |||| |||||||||||||||||  |||  ||| || |||||||||||||||||
Db     995  TTGGCTACGGCAACGGCGACAATATGCCCGATGAGAAGCTGCTGCCGGTTGTTCGACAAA  1054

Qy    1546  TTAATTTAGATAAAGGTTATCATTTGTTTTACATCATAGAGGAGTCATATTCCCT  1605
                  | |||| || ||||  |||| ||  ||| |||| |||||||||||||||| |
Db    1055  TCAATTTGCAGCAGGCAGGCATTTTATCGTGTTGCACCAACGCGGTTCGCACGCCCCAT  1114

Qy    1606  ATGGGGCATTATT  1618
                  |||||||||||
Db    1115  ACGGCGCATTGTT  1127
```

```
RESULT 12
AP005083
LOCUS        AP005083      288108 bp    DNA      linear   BCT 07-APR-2004
DEFINITION   Vibrio parahaemolyticus DNA, chromosome 1, complete sequence,
             11/11.
ACCESSION    AP005083 BA000031
VERSION      AP005083.1  GI:28807856
KEYWORDS     .
SOURCE       Vibrio parahaemolyticus
  ORGANISM   Vibrio parahaemolyticus
             Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
             Vibrionaceae; Vibrio.
REFERENCE    1
  AUTHORS    Nasu,H., Iida,T., Sugahara,T., Yamaichi,Y., Park,K.S., Yokoyama,K.,
             Makino,K., Shinagawa,H. and Honda,T.
  TITLE      A filamentous phage associated with recent pandemic Vibrio
             parahaemolyticus O3:K6 strains
  JOURNAL    J. Clin. Microbiol. 38 (6), 2156-2161 (2000)
  MEDLINE    20295086
   PUBMED    10834969
REFERENCE    2
  .AUTHORS   Makino,K., Oshima,K., Kurokawa,K., Yokoyama,K., Uda,T.,
             Tagomori,K., Iijima,Y., Najima,M., Nakano,M., Yamashita,A.,
             Kubota,Y., Kimura,S., Yasunaga,T., Honda,T., Shinagawa,H.,
             Hattori,M. and Iida,T.
  TITLE      Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
             distinct from that of V cholerae
  JOURNAL    Lancet 361 (9359), 743-749 (2003)
  MEDLINE    22508454
   PUBMED    12620739
REFERENCE    3  (bases 1 to 288108)
  AUTHORS    Hattori,M., Yamashita,A., Oshima,K. and Shiba,T.
  TITLE      Direct Submission
  JOURNAL    Submitted (09-APR-2002) Masahira Hattori, Kitasato Institute for
             Life Sciences, Kitasato University; Kitasato 1-15-1, Sagamihara,
             Kanagawa 228-8555, Japan
             (E-mail:hattori@genome.ls.kitasato-u.ac.jp,
             URL:http://genome.ls.kitasato-u.ac.jp/, Tel:81-42-778-8194,
             Fax:81-42-778-8193)
COMMENT      This work was done in collaboration with Kozo Makino, Ken Kurokawa,
             Katsushi Yokoyama, Takayuki Uda, Kenichi Tagomori, Masatomo Najima,
             Masayuki Nakano, Yoshino Kubota, Shigenobu Kimura, Teruo Yasunaga,
             Takeshi Honda, Hideo Shinagawa, Tetsuya Iida (Osaka University),
             Yoshio Iijima (Kobe Institute of Health), and supported by the
             Research for the Future Program of the Japan Society for the
             Promotion of Science. This clone was isolated from a patient
             presenting with acute gastroenteritis.
FEATURES             Location/Qualifiers
     source          1..288108
                     /organism="Vibrio parahaemolyticus"
                     /mol_type="genomic DNA"
                     /strain="O3:K6"
                     /sub_strain="RIMD 2210633"
                     /db_xref="taxon:670"
                     /chromosome="1"
     gene            91..555
```

```
     CDS             91..555
                     /gene="VP2831"
     gene            91..555
                     /gene="VP2831"
                     /note="similar to GB:AAF95794.1 (AE004331) percent
                     identity 92 in 150 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="protein-transport protein SecB"
                     /protein_id="BAC61094.1"
                     /db_xref="GI:28807857"
                     /translation="MAEAAPQEAQQNFAIQRIFLKDVSFEAPNSPVIFQKEWNPDVKL
                     DLDTQSRELGEGVYEVVLRLTVTVKNEEETAPLCEVQQGGIFTAEQMEAGQLAHCLGA
                     FCPNILFPYARETISSLVVKGTFPQLNLAPVNFDALFMNYLQQQAQQGEAQA"
     gene            752..1789
                     /gene="VP2832"
     CDS             752..1789
                     /gene="VP2832"
                     /note="similar to GB:AAF95792.1 (AE004331) percent
                     identity 86 in 343 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="glycerol-3-phosphate dehydrogenase (NAD+)"
                     /protein_id="BAC61095.1"
                     /db_xref="GI:28807858"
                     /translation="MTQANTNNAYGKDIAMTVIGAGSYGTSLAISLARNGANVVLWGH
                     EPEHMARLEADRANHAFLPGVFDFPESLIIESDLEKAVQASRDLLVVPSHVFGIVLNS
                     CKPFLREDSRICWATKGLEPETGRLLKDVAYDIIGENYSLAVLSGPTFAKELAMGLPT
                     AISVASPDAEFVADLQEKIHCSKTFRVVANNDFIGMQLGGAVKNVIAIGAGMSDGIGF
                     GANARTALITRGLAEMSRLGAALGAKPETFMGMAGLGDLVLTCTDNQSRNRRFGLALG
                     QGQDVDTAQEEIGQVVEGYRNTKEVWMLSQRMGVEMPIVDQIYQVLYQGKDARLAAQD
                     LLARDKKAEGK"
     gene            1868..2689
                     /gene="VP2833"
     CDS             1868..2689
                     /gene="VP2833"
                     /note="similar to GB:AAF95790.1 (AE004331) percent
                     identity 88 in 273 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="serine acetyltransferase"
                     /protein_id="BAC61096.1"
                     /db_xref="GI:28807859"
                     /translation="MKHCEKQKVWNKIVSEAREMSEQEPMLASFYHATIIKHESLCAA
                     LSYILANKLNTASMPAMAVREVVEEAFAADPTITDFRAACDICATVNRDPAVSMYSMPL
                     LYLKGYHALQGYRVANWLWKQGRHALATYLQNOISVACQVDIHPAARIGSGIMLDHAT
                     GIVIGETAVIDNDVSILQDVTLGGTGKECGDRHPKIREGVMIGAGAKILGNIEVGEGA
                     KIGSCSVVLQAVPPHTTVAGVPAKIVGRPKTDKPSLDMDQGFNGKSQSFIHGDGI"
     gene            2755..3861
                     /gene="VP2834"
     CDS             2755..3861
                     /gene="VP2834"
                     /note="similar to GB:AAF93508.1 (AE004122) percent
                     identity 65 in 362 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="NlpD-related protein"
                     /protein_id="BAC61097.1"
                     /db_xref="GI:28807860"
                     /translation="MLRSAILLTCALSATYPLSSFAASQQELKGVSSEINRQKALTS
                     QEKQLNELQKSLKDQELGISKLEREIKQTKADLAQADKNINKLEEKISLQEAQRQAQE
                     EELKQLIQTYYVTERAKANGHLLNQGVEBDRISQYFQHLAKAREVIDAITKTQELA
                     HKNQLELEKEQIETLLKQQSEKRTALASTOSNRKQTLNKIQSSIKNDKRYLAELQRN
                     ETRLKABIAKAAKRNAVPMDGLGKQRGKLPWPLKGSVLLHNFGTRQTGQVNWKGMVLSA
                     NYGQQVKAVYPGTVVFABYLRGYGLVLLLDHGKGDMTLYGYNQALTKKEBGDKVTAGEV
                     IALAGDTGQDRPSLYFEIRRNSEAQNPKSWLKR"
     gene            complement(3938..4894)
                     /gene="VP2835"
     CDS             complement(3938..4894)
                     /gene="VP2835"
```

```
Query Match                8.1%;  Score 134.4;  DB 1;  Length 288108;
Best Local Similarity     61.0%;  Pred. No. 1.4e-12;
Matches  233;  Conservative  1;  Mismatches  147;  Indels  1;  Gaps  1;

Qy      12  CAATAGGSGGATCCACGAGCTTCTATTAGGTATCGTATTGGCTGCGAGGGATATCCAAAG   71
                | ||||||:||||| ||||||||||| ||||||||||| | || |||||||||||||||
Db   82880  CTAAATGGGATCCACGGCCATCTA-TCGGTATCGTGCTAGCAGCCGGTGGCTACCCAGCA   82938

Qy      72  GATTATCGCAAAGGCCATGAAATCACGGATTGCCTAAAAGTGCGGTCAAAAACGAGAAA   131
                ||||||| ||||| ||||||||||||||| ||| ||||| ||||| ||||||| |||||
Db   82939  GCGTACAACAAAGGTGACGTGACGTGATCGTCTGGCCTGCCTCAAGTTGAAATTGAGGGCGAGAAA   82998

Qy     132  GTTTCTTAGCGGTGTCGCAGAACAAGAGGCAAGCTAGTCACAAACGCGCGGTCGTGTA   191
                ||| || ||||| || ||||||||||| ||||| |||||||||||||||| |||||||
Db   82999  GTCTTCCACGCGGGTACAGACAATCAAGATGCGACAATCGTGACACAACAAAAGCGTAAAATTG   83058

Qy     192  CTTTGTGTGACTGCGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTAAAATTG   251
                || |||| || |||| |||||||||||| ||| ||||| |||| |||||| |||||||
Db   83059  TTGTGTGCAACTGCATTGGGCACCAGCGTATCAGAAGCACAACACGCGCATACAGAGCTA   83118

Qy     252  GCTGAGCAAATTCAATGGCTCTGGGCCTTTTTATCGTCGAGACATTGGTTACAGGGCGTG   311
                ||||||| ||||||||| || |||||||| ||||||| ||||||| |||| |||||||||
Db   83119  GCAAAACAGATCAGTTGGGATGGCATGTTCCATCGTAATGACATCGGCTACCGTGCTATC   83178

Qy     312  GAACGAGAACAAGCAAAATAGTTAGAAATCTTGTTGAAATTTAATTAGATAAAAATATTG   371
                ||| ||||| |||| ||||||||| || |||||||||| |||||| || |||||||||||
Db   83179  GCTCGCGAGCAAGAAGAAAATAATAACACTTCAGTTGGCTAAAAAGACAAAAGGCGCATTC   83238

Qy     372  TACAGGTAGAATTGTATTTC   393
                |||||||||||||||||||||
Db   83239  AGCGCTTTTTTATTGCTCGTC   83260


RESULT 13
AE017153/c
LOCUS       AE017153              300521 bp    DNA     linear   BCT 15-AUG-2003
DEFINITION  Haemophilus ducreyi strain 35000HP section 3 of 6 of the complete
            genome.
ACCESSION   AE017153 AB017143
VERSION     AE017153.1  GI:33148155
KEYWORDS    .
SOURCE      Haemophilus ducreyi 35000HP
  ORGANISM  Haemophilus ducreyi 35000HP
            Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
            Pasteurellaceae; Haemophilus.
REFERENCE   1  (bases 1 to 300521)
  AUTHORS   Munson,R.S. Jr., Ray,W.C., Mahairas,G., Sabo,P., Mungur,R.,
            Johnson,L., Nguyen,D., Wang,J., Forst,C. and Hood,L.
  TITLE     The Complete Genome Sequence of Haemophilus ducreyi
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 300521)
  AUTHORS   Munson,R.S. Jr., Ray,W.C., Mahairas,G., Sabo,P., Mungur,R.,
            Johnson,L., Nguyen,D., Wang,J., Forst,C. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUN-2003) Pediatrics, Columbus Children's Research
            Institute and The Ohio State University, 700 Children's Drive,
            Columbus, OH 43205, USA
FEATURES             Location/Qualifiers
     source          1..300521
                     /organism="Haemophilus ducreyi 35000HP"
                     /mol_type="genomic DNA"
                     /strain="35000HP"
                     /db_xref="taxon:233412"
     gene            90..614
                     /locus_tag="HD0771"
     CDS             90..614
                     /locus_tag="HD0771"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAP95676.1"
                     /db_xref="GI:33148156"
```

```
                     /note="similar to GB:AAF93507.1 (AE004122) percent
                     identity 56 in 318 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAC61098.1"
                     /db_xref="GI:28807861"
                     /translation="MNYDIFNGDADGIIALLQLRLADPIDSQLITGVKRDIKLVEKVD
                     VQAGDELTVDISMEKNMAGLEQALAQGAHVFYADHHKAGDIPQHGNLDAHIDLDANM
                     CTALIVDKLLEGRFHTWAITAAYGDNLIAKADVLADQAGLNNEQKAQLKELGTLINYN
                     GYGSKVDDLHFHPADLYRALVQYISPFEVIEDKASTYYQLQSAYQQDMDAAQAVPATH
                     ESDTLKLFELPNTAASRRIXGVYGNWLANQNPDSAHAVLTENADGTYTVSLRAPLNNK
                     QGAVAVCGQFPTGGGREAAAGINALRKEDVNAFIDAVETYYA"
     gene            complement(5042..5677)
                     /gene="VP2836"
     CDS             complement(5042..5677)
                     /gene="VP2836"
                     /note="similar to GB:AAF93506.1 (AE004122) percent
                     identity 74 in 209 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="transcriptional regulator, TetR family"
                     /protein_id="BAC61099.1"
                     /db_xref="GI:28807862"
                     /translation="MKTRDKIVYAALELENQHGRNITTNHIADHIEISPGNLYYHFR
                     NKQEIVREIFALYSAELLERFTPIQGSQESLTMLKSYLDSIFTLMWKYRFFYANLPEI
                     LSRDEQLHEQYIDVQEKLQANLIAIMQEFVSLKLDVNEQQLKSLVCTLHLIACSWLA
                     YQSAMASKTSITEQMVKQGMLQMLNVVKPVATEQGLEQLQLLEEAVSTLQG"
     gene            complement(5699..6073)
                     /gene="VP2837"
     CDS             complement(5699..6073)
                     /gene="VP2837"
                     /note="unknown"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAC61100.1"
                     /db_xref="GI:28807863"
                     /translation="MNLTAARFVILSYAILIAFLSLSSGELQDWGNIAYLDKMQHAMA
                     YSGPAFLGCFCVWRRQRWAVAMGILLFSAVIEWLQGYVGRQTSWLDLVANSTGILTG
                     LLCYRIWSLYRPRLVKSLYSIR"
     gene            complement(6083..7054)
                     /gene="VP2838"
     CDS             complement(6083..7054)
                     /gene="VP2838"
                     /note="similar to GB:AAF95796.1 (AE004331) percent
                     identity 83 in 321 aa"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="lysyl-tRNA synthetase-related protein"
                     /protein_id="BAC61101.1"
                     /db_xref="GI:28807864"
                     /translation="MQTNWQPTASIEQLRQRATLIAAIRQFFAARQVMEVDTPAMSHA
                     TVTDIHLHTFQTEFVGPGYADGSKLFFMTSPEFHMKRLLAAGSSCIYQINKAFRNEEN
                     GRYHPEFTMLEWYRVGFDHHKLMDEMDDLLQLVLKCGAAQRMTYQQAFIDVLGVCPL
                     EGSMTELKAAASKLGLSDIAEPEEDRDTLLQLLFSVGVENKIGQDVPAFVYDFPASQA
                     ALAKINPQDHRVADRFEVYFKGIELANGFWGIELDNPKEQLARFEQDNAKRIEMGLKPQP
                     IDYHLISALEAGLPDCAGVALGIDRLIMLALGCDHIDQVTAFFPIA"
     gene            7158..7280
                     /gene="VP2839"
     CDS             7158..7280
                     /gene="VP2839"
                     /note="unknown"
                     /codon_start=1
                     /evidence=not_experimental
                     /transl_table=11
                     /product="hypothetical protein"
                     /protein_id="BAC61102.1"
                     /db_xref="GI:28807865"
```

```
                    /translation="MQKVKLPLTIDPFKDAQRRVDVVGYFSSKQLVRLGESVCNVLND
                    AQVTLSLIYVDPQRLSVLKGSAQVEVELCCQRCGNSFTQVIDCSFLFSPVRNLDQAGDL
                    PEIYEPEVNAFGEVNLLDMIRDEFILALPLVPMHSEAHCEVSVSEQVFGELPAELAK
                    KPNPFAVLANLKKN"
gene                633..803
                    /gene="rpmF"
                    /locus_tag="HD0772"
CDS                 633..803
                    /gene="rpmF"
                    /locus_tag="HD0772"
                    /codon_start=1
                    /transl_table=11
                    /product="50S ribosomal protein L32"
                    /protein_id="AAP95677.1"
                    /db_xref="GI:33148157"
                    /translation="MAVQQNKKSRSRRDMRRSHDALTTAAVSVDKTTGETHLRHHVTA
                    DGYYRGRKVINK"
gene                825..1856
                    /gene="plsX"
                    /locus_tag="HD0773"
CDS                 825..1856
                    /gene="plsX"
                    /locus_tag="HD0773"
                    /codon_start=1
                    /transl_table=11
                    /product="fatty acid/phospholipid synthesis protein"
                    /protein_id="AAP95678.1"
                    /db_xref="GI:33148158"
                    /translation="MNRITLALDVMGGDFGPRVTIPALSLALAQQPMLSFVLFGDQYH
                    AIPLLHSLPAEQQORIRFVHTSRIVDPHVPVRQALRQSKGSSMRLAIEAVVAGEAQAC
                    VSGGNTGVLMGLAKWFVEPLVGIDRPALASLIPALNGQASVMLDLGANVEADSKLLLQ
                    FAEMGNIFAQSLSLSLAQPRLALLNIGTBEVVGKRYAIQQAHHVLKNRADLHYIGFIESD
                    KLMNHLADVIVCDGFCGNIALKAVEGVASHFLALFKQPKRSFFVDWLSSWWGRLLYHR
                    VYRALQRVNPERHNGATILGLSSVVVKSHGSAGVNAYFYAINHAIEQIQNRIPVKISA
                    GLQKLHQNL"
gene                1868..2818
                    /gene="fabH"
                    /locus_tag="HD0774"
CDS                 1868..2818
                    /gene="fabH"
                    /locus_tag="HD0774"
                    /codon_start=1
                    /transl_table=11
                    /product="3-oxoacyl-(acyl-carrier-protein) synthase III"
                    /protein_id="AAP95679.1"
                    /db_xref="GI:33148159"
                    /translation="MYSKILATGSYLPKQVRSNADLEKMVDTTDEWIFTRSGMKERRI
                    AAADETVATMGAAAAHKACEMAAIDVNEIELIVVATTHSHAYPSAACQIQGMLAIED
                    AIAFDVAAACTGFIYALSVADQPIRAGKVKTAFVGVADVNSRVLDEADRGTVVLFGDG
                    AGAVILQASEQAGILSTHLHSSADTDNMLALPIQERGNAQSGFIQMQGNATFKLAVSR
                    LANVVEETLQANNLQKADLDWLVPHQANIRIIAATAKKLEMEMSQVVLTVEKYGNNSA
                    ATVPVALDEAVRDGRIQRGQLLLLEARFGGGWTWGSGLVRF"
gene                2828..3577
                    /gene="HD0775"
                    /locus_tag="HD0775"
CDS                 2828..3577
                    /gene="HD0775"
                    /locus_tag="HD0775"
                    /codon_start=1
                    /transl_table=11
                    /product="probable methyl transferase"
                    /protein_id="AAP95680.1"
                    /db_xref="GI:33148160"
                    /translation="MSEKYTFADQKRRKTVEQAEFTEDGRYLRKVRSFVLRTGRLSDFQ
                    RDMMNDHWAKFGLDYQKSPFDFVEIFGMGRSLVEMAAQNPDRNYIG
                    IEVHTPGVGACIAYAVEKQVDNLRVICHDATEILRDAVADGSLAGLQLYFPDPWQKSK
                    HHKRRIVQTEFVMRLLTKLTAGGFIHFATDWQNYAEHMLAVLRQFSPALVNTSPTGDF
                    IDRPAFRPLTKPEARGQRLGHAVWDLYFVKK"
gene                3658..3999
                    /locus_tag="HD0776"
CDS                 3658..3999
                    /locus_tag="HD0776"
                    /codon_start=1
                    /transl_table=11
                    /product="conserved hypothetical protein"
```

```
                    /protein_id="AAP95681.1"
                    /db_xref="GI:33148161"
                    /translation="MAIQRNARQRKKMHLABFQELGFLVKFQFAENTGIDQIDAAVDR
                    FINEVIKPNGLAYEGSGYLHWEGLVCLEKLGKCEASHQQIVKQWLESNGLSQVEISEL
                    FDIWWDYPTEV"
gene                4138..4671
                    /locus_tag="HD0777"
CDS                 4138..4671
                    /locus_tag="HD0777"
                    /codon_start=1
                    /transl_table=11
                    /product="conserved hypothetical protein"
                    /protein_id="AAP95682.1"
                    /db_xref="GI:33148162"
                    /translation="MQFLAPIISIVVGFLSFVLVLRTWMQFCQVSAYLPLSQSLLRLT
                    APMVNPVSKFIPTVKNINFAALLIALALLAVEKGVLGLPLTMAVLVGLLGIIKMFGQI
                    LLNNLFYSVFGAFWAVA"
gene                4671..4964
                    /locus_tag="HD0778"
CDS                 4671..4964
                    /locus_tag="HD0778"
                    /codon_start=1
                    /transl_table=11
                    /product="conserved hypothetical protein"
                    /protein_id="AAP95683.1"
                    /db_xref="GI:33148163"
                    /translation="MKAVEFVENPYGIRLRIFLQPKASRDQIVGLHDNELKVAITAPP
                    VDGAANAYLLKYLSKLFKVPKSSIVLEKGELQRHKQLFVPAPKLLPKEIEQWL"
gene                complement(5071..6473)
                    /locus_tag="HD0779"
                    /note="product: pseudogene for cytochrome c peroxidase,
                    authentic frameshift.
                    This region contains an authentic frameshift or in-frame
                    stop in the coding sequence and is not the result of a
                    sequencing error."
                    /pseudo
misc_feature        complement(5994..6473)
                    /note="cytochrome c peroxidase, N-terminal fragment
                    HDp0779A"
gene                complement(6638..6961)
                    /gene="phnA"
                    /locus_tag="HD0781"
CDS                 complement(6638..6961)
                    /gene="phnA"
                    /locus_tag="HD0781"
                    /codon_start=1
                    /transl_table=11
                    /product="alkylphosphonate uptake protein"
                    /protein_id="AAP95684.1"
                    /db_xref="GI:33148164"
                    /translation="MTYPICPTCKSENTYHDTIQFVCPDCAYEWTGEEINTADEDQLI
                    VKDSNGNLLADGSVLLIKDLKLKGSSEVLKKGTKFPKGIRLVDGDHNVDCGKIMLKSE
                    FLKKA"
gene                complement(7078..8385)
                    /gene="oadB"
                    /locus_tag="HD0782"
CDS                 complement(7078..8385)
                    /gene="oadB"
                    /locus_tag="HD0782"
                    /codon_start=1
                    /transl_table=11
                    /product="oxaloacetate decarboxylase beta chain"
                    /protein_id="AAP95685.1"
                    /db_xref="GI:33148165"
                    /translation="MESIIALFQGMGIMHITLGQVIMIIVSLILLWLAIARRFRPLLL
                    LPIGFGGLLSNIPEAGLAMTALDHLLHYAHTEQLTIIAEKVNSSSLDVHTIKQAIAFA
                    PPSVQNELEVIASDLGYNAGIIALFYKVAIGYGVAPLIIFMGVGAMTDFGPLLANPRT
                    LLLGAAAQFGIFTTVLGALGLNWLGIIDFSLPQAAAIGIIGGADGPTAYILASKLAPE
                    LLGAIAVAAYSYMALVPLIQPPIMKALTTEQERKIRMVQLRTVSAREKIVFPIVLLLL
                    VALLLPDAAPLLGMFCFGNLMRASGVVERLNETAQNALINIVTIFLGLSVGAKLVADK
                    FLQPQTLGILLLGIVAFAIGTASGVIMAKIMNAFSKHKINPLIGSAGSVAPMAARVS
                    NKIGLEADKQNFLLMHAMGPNVAGVISSAIAAGIMLKYISTMI"
```

```
gene            complement(8397..10202)
                /gene="oadA"
                /locus_tag="HD0783"
CDS             complement(8397..10202)
                /gene="oadA"
                /locus_tag="HD0783"
                /codon_start=1
                /transl_table=11
                /product="oxaloacetate decarboxylase alpha chain"
                /protein_id="AAP95686.1"
                /db_xref="GI:33148166"
                /translation="MAIQNKKIAITDVVLRDAHQSLFATRLRLDDMLPIATELDKVGY
                WSLETWGGATFDSCIRFLGEDPWVRLRELKKAIPNTPLQMLLRGQNLLGYRHYADDVV
                DRFVDRCIANGMDIPRIFDALNDPRNMQAALQAVKKHGGHAQGTLSYTTSPVHTLQTW
                LDITEQLLEIGIDSLVIKDMSGILTPMAAYELVSEIKKRYEIELHLHCHSTTGMAEMA
                LLKAIEAGIDGIDTSISSMSGTYGHPATESMVATLQGTAYDTGLNIHQLEKIAAYFRQ
                VRKKYAKFEGQLRGIDSRILVAQVPGGMLTNLESQLKQQNATDKLDLVLQEIPKVRED
```

Query Match   7.3%; Score 121.6; DB 1; Length 300521;
Best Local Similarity 67.1%; Pred. No. 1.6e-10;
       Matches 188;   Conservative   0; Mismatches   89; Indels   3; Gaps   1;

```
Qy       79 GCAAAGGCGATGAAATCAGCGGATTGCCTAAAAGTGCGGTCAAAAACAGAAAGTTTTCT 138
            | ||  ||| ||| |||||| |||||||  ||| |||| |   | ||| |||  || ||
Db   226828 GGAAGGGCGATGAAATTACCGCTATTCCTACTACACTCCACCAC---AAGGTTTTTC 226772

Qy      139 TAGCGGGTGTCGCAGAACAAGAGGCAAGCTAGTCACAAACGGCGGTCGTGTACTTTGTG 198
            |||| ||||| |||||| ||||| |||  || ||||||  | ||| ||||||||| |||
Db   226771 TAGCGGGTGTCGAAGAAAGGCGGGTAAATTAGTTGTTCACTAATGGCGGTAGAGTATTATGTG 226712

Qy      199 TGACTGCGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAAATTGGCTGAGC 258
            ||  || | || ||   ||| ||  |||| |   |||| |  || |||  |||| || |
Db   226711 CAACGGCTTTAGGTAATTCAGTGTATGAAGCTCAACAACAAGCATTGAAATTGGCTGAAC 226652

Qy      259 AAATTCAATGGTCTGGGCGTTTTATCGTCGAGACATTGGTTACAGGGCTGTGGAACGAG 318
            ||| |   ||| || | |||||  | || |   |||||| ||    | |||||||| ||
Db   226651 AAATTAATTGGCAAGACGCGTTTCTATCTGCGGATATTGGTTTATCTGCCGTACAGCGCG 226592

Qy      319 AACAAGCAAAATAGTTAGAAATCTTGTTGAATTAATTAG 358
            |||| |||||| |  |||| | |||| |
Db   226591 AAAAAGCATAATTAATGGCGATGTTATCCGCTTAAATAAG 226552
```

RESULT 14
A96274

LOCUS   A96274   1275 bp   DNA   linear   PAT 07-SEP-2000
DEFINITION Sequence 307 from Patent WO9924578.
ACCESSION   A96274
VERSION   A96274.1 GI:6780029
KEYWORDS   .
SOURCE   unidentified
  ORGANISM unidentified
            unclassified.
REFERENCE   1
  AUTHORS   Pizza,M., Scarlato,V., Rappuoli,R., Grandi,G. and Masignani,V.
  TITLE   Neisserial antigens
  JOURNAL   Patent: WO 9924578-A 307 20-MAY-1999;
            PIZZA MARIAGRAZIA (IT); SCARLATO VINCENZO (IT); RAPPUOLI RINO (IT);
            CHIRON SPA (IT); GRANDI GUIDO (IT); MASIGNANI VEGA (IT)
FEATURES   Location/Qualifiers
     source   1..1275
                     /organism="unidentified"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32644"
ORIGIN

Query Match   7.3%; Score 120.8; DB 6; Length 1275;
Best Local Similarity 50.0%; Pred. No. 5.6e-10;
       Matches 384;   Conservative   4; Mismatches 362; Indels 18; Gaps   3;

```
Qy      589 CTGAAAGTTATTTGTCTTGTTGTTGTTGTTGTTTTTCAAAATATAGAA 648
            | ||  ||| || |||||||||| ||||||||||||| |
Db      110 CAGAAACGTTTGCGCTGACATTTGTGATTGCTGCGCTGTATCTGTTTGCGCGTTATAAGG 169
```

RESULT 15
NMA2Z2491/c

LOCUS   NMA2Z2491   349061 bp   DNA   linear   BCT 02-SEP-2002
DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome;
            segment 2/7.
ACCESSION   AL162753 AL157959
VERSION   AL162753.2 GI:7379120
KEYWORDS   .
SOURCE   Neisseria meningitidis Z2491
  ORGANISM Neisseria meningitidis Z2491
            Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
            Neisseriaceae; Neisseria.
REFERENCE   1 (bases 1 to 349061)
  AUTHORS   Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C.,
            Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T.,
            Davies,R.M., Davis,P., Devlin,K., Feltwell,T., Hamlin,N.,
            Holroyd,S., Jagels,K., Leather,S., Moule,S., Mungall,K.,
            Quail,M.A., Rajandream,M.A., Rutherford,K.M., Simmonds,M.,
            Skelton,J., Whitehead,S., Spratt,B.G. and Barrell,B.G.
  TITLE   Complete DNA sequence of a serogroup A strain of Neisseria
            meningitidis Z2491
  JOURNAL   Nature 404 (6777), 502-506 (2000)
   MEDLINE   20222556
```

```
Qy      649 TTACAAGAAATATTAGTTGGGCTTTATTTGCAATAAGTAATATGTTGTTGTTAATAATGTACATT 708
            |||||||||| || ||| || |||||||| |||| ||| || || |||||||
Db      170 CAACGCGTTTGTTGATTGCGGTGTTTTTCGCGTTGATCAGCATTATTGCCAACAATGTGCATT 229

Qy      709 ATGCAGTATACCAATCTTGGATTGGACCTGTGTAATTACTCACTGCATTTAAAGAAATTA 768
            ||||||| |   || |||| ||| || |||  |||| ||| ||||||||   |||| |||
Db      230 ACGCCGGTTATCAAAGCTGGATAACGGCACATTAATTATTGGCTGATGCTGAAAGAGATTA 289

Qy      769 ATGAGATAACAAATGCTGGCTTAACAATGATAGATAAATTCATTATATCCATTGTTATTTG 828
            || |   || |  ||| |||  | |  || |||| | ||||| ||| ||| |||| ||
Db      290 CCGAAGTTGGCGCGCCAGGGGCGTCGATGTTGGATAAGTTGTGGCCTGCCGTTGTGGG 349

Qy      829 GTTTATTTGAAGTTGCTGTGTTTTAAGTTAAGTTCATAAAAGAAAGTATATAAAC 888
            | ||| |||| ||||| ||| || ||||| || |||   || || ||| ||||
Db      350 GCGTGTTGGAAGTCATCGTGTTGTTTGCAGCCTTGCCAAGTTCCGCGTAAGACGCAT--- 405

Qy      889 TTTCTTGGATTTTGACTTATTTTATGCTGTGATGATGGTATGTTTTGTTGTCGAGCGT 948
            ||||| ||  |||
Db      406 -------TTTCTGCCGATATACTGTTTGCCTTCCTAATGCTGTGATGATTTTTCGTGCGTT 457

Qy      949 ATACAACAAAATCCCATGAGCGTTTTATTCACCTAACACTGTTTATTCTCGATTAAAAT 1008
            | ||| |  | |||| | ||| |||| || |||||||||||  | |||||||| | |||
Db      458 CGTTCGACACGAAACAAGAACACGGTATTTCGCCCAAACCGACATACAGCGCCATCAAAG 517

Qy     1009 CCAATTATTATCGTTGGTATTTTATAGGACGAATTGTTCCTTATGAGATATTTTCTT 1068
            |||||||||| |||||| ||| |  || ||||| |||| | || |||| ||| | |||
Db      518 CCAATTATTTCAGCTTCGGTTATTTTGTCGGACGCGTGTTGCCGTATCAGTTGTTTGATT 577

Qy     1069 TATCTAATATTCCTCTTTATCATAAATCTAAGCCACTGCTTCCAAGCAGAATCGGGGCTCTCCGVAAAATT 1128
            |||  ||| | |  ||||| ||| |   | ||| | |||  |||||| |||   ||||| ||
Db      578 TAAGCAAGATTCCTGTTGTTCAAACAGCAGCAGCCTGCTCCAAGCAGCATTGGGGTAAATTGGCAGTATTC 637

Qy     1129 MARGAATATAATTTTAATTAATGGGGAAGTGCGACCTCAAGTCATTTTAGTGCTTTTG 1188
            || |  |     |||||||||  |  | || ||| | |  ||||| | ||||| |||
Db      638 AAAATCGTCCTGA---TTATGGCGGCAAAGCGAAAGGCGCGCATTTGAAATTGTTTG 694

Qy     1189 GTTACGGGAGAAAAACATCTCCTTTTT---AGATAGCTTAAAATATAAAATCAGGAGCTC 1245
            ||  || | |   | |||| || |||    | |||| || ||  ||| || ||||| ||
Db      695 GCTACGGGCGGCGAAACTTCGCCGTTTTGACCCAGCTTCGCAAGCCGATTTTAAGCCGA 754

Qy     1246 TTGTTCGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTTCTTTACCAATGTTTTTTA 1305
            ||||| |||| |   |||| |  | |  ||| ||||||  | |||| |  ||||| |||
Db      755 TTGTGAAACAAAGTTATTTCCCGAGCGCCCAGGCTTTGAACAGATAATCAGCGCCGGCGGTA 814

Qy     1306 ATGCAATYCCTTAYCCAAATGAATACAACAGATAGCTAAAGGAGATA 1353
            | |||||||  ||| |  ||| || || |||| |  || || || |
Db      815 ACGTCATACCGGATGCCAACGCGCTTGGAACAAAATCAGCGGCGGCGGCGATA 862
```

```
PUBMED      10761919
REFERENCE   2  (bases 1 to 349061)
AUTHORS     Parkhill,J.
TITLE       Direct Submission
JOURNAL     Submitted (30-MAR-2000) Submitted on behalf of the Neisseria
            sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT     Notes:
            Details of N. meningitidis sequencing at the Sanger Centre are
            available on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
FEATURES             Location/Qualifiers
     source          1..349061
                     /organism="Neisseria meningitidis Z2491"
                     /mol_type="genomic DNA"
                     /strain="Z2491"
                     /db_xref="taxon:122587"
                     /note="serogroup: A"
     repeat_unit     complement(24..206)
                     /note="ATR repeat; hmmfs hit to HMM ATR (1 - 183), score:
                     310.39"
                     /label=ATR
     RBS             209..212
     gene            220..681
                     /gene="NMA0368"
     CDS             220..681
                     /gene="NMA0368"
                     /note="NMA0368, probable integral membrane protein, len:
                     153 aa; contains four probable transmembrane domains"
                     /codon_start=1
                     /transl_table=11
                     /product="putative integral membrane protein"
                     /protein_id="CAB83669.1"
                     /db_xref="GI:7379121"
                     /db_xref="TrEMBL:Q9JWH7"
                     /translation="MQEQNRKSSFPIVMLLVSVALWIASLSNVAFYLGNHGSMEGLTV
                     LILGSIFASLDIRYCAVYANYVWLAAIVLLALRKKVVPVHAAFWGLALVAFSVKAVYV
                     DEAGNTSDIVRYGAGFYLWYAAFAVASIGTFAGKNKERKAASAADGTKNDV"
     gene            671..1492
                     /gene="hemK"
     CDS             671..1492
                     /gene="hemK"
                     /note="NMA0369, hemK, HemK protein, len: 273 aa; similar
                     to e.g. HEMK_ECOLI P37186 HEMK protein (277 aa), fasta
                     scores; E(): 0, 42.3% identity in 279 aa overlap. Contains
                     PS00092 N-6 Adenine-specific DNA methylases signature"
                     /codon_start=1
                     /transl_table=11
                     /product="HemK protein"
                     /protein_id="CAB83670.1"
                     /db_xref="GI:7379122"
                     /db_xref="GOA:Q9JWH6"
                     /db_xref="TrEMBL:Q9JWH6"
                     /translation="MTFDKWLGLSKLPKNEARMLLQYVSEYTRVQLLTRGGEMPDEV
                     RQRADRLAQRRLNGEPVAYILGAREFYGRRFTVNPSVLIPRPETEHLVEAVLARLPEN
                     GRVWDLGTGSGAVAVTVALERPDAFVRASDISPPALETARKNAADLGARVEFAYGSWF
                     DTDMPSEGKWDIIVSNPPYIENGDKHLSQGDLRFEPQIALTDFSDGLSCIRTLAQGAP
                     DRLAEBGFPLLLEHGFDQGAAVRGVLAENGFSGGVETLPDLAGLDRVTLGKYMKHLK"
     misc_feature    complement(1279..1288)
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
     misc_feature    1162..1171
                     /gene="hemK"
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
     misc_feature    1187..1207
                     /gene="hemK"
                     /note="PS00092 N-6 Adenine-specific DNA methylases
                     signature"
     misc_feature    complement(831..840)
                     /gene="hemK"
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS

     RBS             1565..1568
     gene            1578..2966
                     /gene="NMA0370"
     CDS             1578..2966
                     /gene="NMA0370"
                     /note="NMA0370, probable integral membrane protein, len:
                     462 aa; similar to hypothetical proteins e.g. Y325_HAEIN
                     P44640 hypothetical protein HI0325 (450 aa), fasta scores;
                     E(): 0, 49.8% identity in 464 aa overlap"
                     /codon_start=1
                     /transl_table=11
                     /product="putative integral membrane protein"
                     /protein_id="CAB83671.1"
                     /db_xref="GI:7379123"
                     /db_xref="GOA:Q9JWH5"
                     /db_xref="TrEMBL:Q9JWH5"
                     /translation="MNAVVVAVIVMLVLSLSRVHVVLSLTIGAFVGGAVAGMPLQNIA
                     DAAGQVSQAGIIPVFNKGLEGGAKIALSVAMLGFAMAITHSGLPQQLAGAVVRKLNR
                     GGMPDSVRSGEGVVKWLLLSILLVGMWSQNIIPIHIAFIPMIVPPLLLVFNRLKIDR
                     RLIACVITFGLVTTYMFLPYGFGAIFLNEILLGNIHSAAPQLDVKNINVMAAMAIPAL
                     GMLAGLLLAFVHYRKPRLYQSNNADTAGNADAANRPQPSAYRSLAAAVAIAVCFAIQL
                     MYEDSLVLGAMLGFAVFMMLGVINRDKANDVFGEGIKMMAMVGFIMIAAQGFAAVMNA
                     TGHIQPLVESSMAIFGNSKGMAALAMLVVGLLVTMGIGSSFSTLPIIAAIYVPLCVGL
                     GFSPLATVAIVGTAGALGDAGSPASDSTLGPTMGLNADGQHDHIRDSVIPTFIHYNIP
                     LLIAGWIAAMVL"
     gene            2967..3191
                     /gene="slyX"
     CDS             2967..3191
                     /gene="slyX"
                     /note="NMA0371, slyX, SLYX protein homolog, len: 74 aa;
                     simlar to SLYX_HAEIN P44759 SLYX protein homolog (73 aa),
                     fasta scores; E(): 0.0017, 33.8% identity in 74 aa
                     overlap, and SLYX_ECOLI P30857 SLYX protein (72 aa), fasta
                     scores; E(): 0.24, 32.4% identity in 68 aa overlap"
                     /codon_start=1
                     /transl_table=11
                     /product="SLYX protein homolog"
                     /protein_id="CAB83672.1"
                     /db_xref="GI:7379124"
                     /db_xref="TrEMBL:Q9JWH4"
                     /translation="MDAVQEFEHRITELEIQSALQEDVIAGLNAMVAELRQTLDIQQA
                     QLRLLYQKMQDRNPDAQEFYSLRDEIPPHY"
     gene            complement(3271..3423)
                     /gene="NMA0372"
     CDS             complement(3271..3423)
                     /gene="NMA0372"
                     /note="NMA0372, unknown, questionable CDS, len: 50 aa"
                     /codon_start=1
                     /transl_table=11
                     /product="very hypothetical protein NMA0372"
                     /protein_id="CAB83673.1"
                     /db_xref="GI:7379125"
                     /db_xref="TrEMBL:Q9JWH3"
                     /translation="MFGRLSPLGSDGFFLAVYEYPVGFICLQGKTAYFRSGGKRFHRI
                     KGILSD"
     misc_feature    3562..3571
                     /note="Core DNA uptake sequence: gccgtctgaa"
                     /label=DUS
     gene            complement(3586..4356)
                     /gene="thiF"
     CDS             complement(3586..4356)
                     /gene="thiF"
                     /note="NMA0373, thiF, probable ThiF protein, len: 256 aa;
                     similar to e.g. THIF_ECOLI P30138 THIF protein (251 aa),
                     fasta scores; E(): 0, 43.1% identity in 246 aa overlap,
                     and MOEB_ECOLI P12282 molybdopterin biosynthesis MOEB
                     protein. (249 aa), fasta scores; E(): 0, 43.9% identity in
                     244 aa overlap (note that N.m. does not have orthologs of
                     any other molybdopterin biosynthesis proteins). Contains
                     Pfam match to entry PF00899 ThiF_family, ThiF family"
                     /codon_start=1
                     /transl_table=11
                     /product="ThiF protein"
```

```
Qy 1129 MARGAATATAATTTTAATTAATGGGGAAAGTGCGACCTCAAGTCATTTAGTGCTTTTG 1188
         :|:  ||     |||||||  | ||||||  || ||  |  ||||  ||||  ||||
Db 65145 AAAATATCGTCCTGA---TTATGGGCGAAAGCGAAAGCGCGGCGCATTTGAAATTGTTTG 65089

Qy 1189 GTTACGGGAGAAAACATCTCCTTTTT---AGATAGCTTAAAATATAAATCAGGAGCTC 1245
         | ||||||  | |||| || || |||      ||||||  |  ||||||||||| ||
Db 65088 GCTACGGGCGCGAAACTTCGCCGTTTTTGACCCAGCTTTTCGGAACGCCGATTTAAGCCGA 65029

Qy 1246 TTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTCTTTACCAATGTTTTTA 1305
         |||| ||| ||| ||||||  | |   | || |||||| ||||| ||| ||| ||||
Db 65028 TTGTGAAACAAAGTTATTCCCGCAGGCTTTATGACGGCAGTATCCCTGCCCAGTTTCTTTA 64969

Qy 1306 ATGCAATYCCTTAYCCAAATGGAATACAACAGATAGCTAAAGGAGATA 1353
         | || ||| |  |   || | ||| |||||||||| | ||| | |||
Db 64968 ACGTCATACCGCATGCCAACGGCTTGGAACAAATCAGCGGCGGCGGCGATA 64921
```

Search completed: December 23, 2004, 18:16:17
Job time : 7041 secs

```
                    /protein_id="CAB83674.1"
                    /db_xref="GI:7379126"
                    /db_xref="GOA:Q9JWH2"
                    /db_xref="TrEMBL:Q9JWH2"
                    /translation="MTTTEHDNDDAFLLRYSRHILLDEIGIEGQQKLSAHILVVGCG
                    GLGAAALPYLAASGIGTLTIADSDTVELHNLQRQVAPDEGDVGKLKTEALADRLRHIN
                    HTVDVRTINEKLDGCRLTGLVQAADIVLDCCDNYATRQAVNRACVQAKTPLVSGAAVR
                    FEGQLAVYRPDLPDSPCYACLFDGGSASDGICSLFGVFSPLVGIIGSTQAEEALKILL
                    DAGEPSHGRLAVYRALEGGWQYFDLPRNPECPVCGAER"
     misc_feature   3787..3796
                    /note="Core DNA uptake sequence: gccgtctgaa"
                    /label=DUS
     misc_feature   complement(3847..4257)
                    /gene="thiF"
                    /note="Pfam match to entry PF00899 ThiF_family, ThiF
                    family, score 186.60, E-value 4.1e-52"
     misc_feature   complement(4418..4427)
                    /note="Core DNA uptake sequence: gccgtctgaa"
                    /label=DUS
     gene           4471..7224
                    /gene="ppc"
     CDS            4471..7224
                    /gene="ppc"
                    /EC_number="4.1.1.31"
                    /note="NMA0374, ppc, phosphoenolpyruvate carboxylase, len:
                    917 aa; simlar to many e.g. CAPP_RHOPA O32483
                    phosphoenolpyruvate carboxylase (EC 4.1.1.31) (936 aa),
                    fasta scores; E(): 0, 43.3% identity in 928 aa overlap.
                    Contains 2x Pfam match to entry PF00311 PEPcase,
                    Phosphoenolpyruvate carboxylase, PS00017 ATP/GTP-binding
                    site motif A (P-loop), and PS00393 Phosphoenolpyruvate
                    carboxylase active site 2"
                    /codon_start=1


Query Match         7.3%; Score 120.8; DB 1; Length 349061;
Best Local Similarity   50.0%; Pred. No. 2.1e-10;
Matches  384; Conservative  4; Mismatches 362; Indels  18; Gaps  3;

Qy    589 CTGAAAGTTATTTGTCTGTTGTTGTTGTTGTTGTTGTTGTTTTCAAAATATAGAA 648
          ||||| ||  | |||||||||||||||||||||||||||||
Db  65673 CAGAAACGTTGCGCTGACATTTGTGATTGCTGCGCTGTATCTGTTTGCGCGTTATAAGG 65614

Qy    649 TTACAAGAATATTAGTTGGCGCCTTATTTGCAATAAGTATTGTTGTTAATAATGTACATT 708
Db  65613 CAACGCGTTTGTTGATTGCGGGTGTTTTTCGCGTTGCCATTATTGCCAACAATGTGCATT 65554

Qy    709 ATGCAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTTGCATTTAAAGAAATTA 768
Db  65553 ACGCGGTTTATCAAAGCTGGATAACGCGGCATTAATTATTGGCTGATGCTGAAAGAGATTA 65494

Qy    769 ATGAGATAACAAATGCTGGCTTAACAATGATAGATAAATTCATATATCCATTGTTATTTG 828
Db  65493 CCGAAGTTGGCGGCGCCAGGGGCGTCGATGTTGGATAAGTTCTGGGCCTGCCTGCGTGGG 65434

Qy    829 GTTTATTTGAAGTTGCTGTGTGTTTTTAAGTTTAAGTTTCATAAAAAGAAAAGTATATAAAC 888
Db  65433 GCGTGTTGGAGTCATCGTTGTTTGTTTGCAGCCTTGCCAAGTTCCGCCTAGAACGCAT---- 65378

Qy    889 TTTCTTGGATTTTGACTTTATTTTATTTTTATGCGTGATGATGTATGTTTTGTTCGAGCGT 948
Db  65377 --------TTTTCTCGCCGATATACTGTTTGCCTTCGCTTCACCTAACACTGTTTATTCTCGATTAAAAC 65326

Qy    949 ATACAACAGAAATCCCATGAGCGTTTTATTTCACCTAACACTGTTTATTCTCGATTAAAAT 1008
Db  65325 CGTTCGACACGAAACAAGAACACGGTATTTCGCCCAAACGACATACAGCCGCATCAAAG 65266

Qy   1009 CCAATTATTTATCGTTGGGTTATTTTATAGGACGAATTGTTCCTTATGAGATATTTTCTT 1068
Db  65265 CCAATTATTTCAGCTTCGGTTATTTTGTCGACGACGCCTGTTGCCGTATCAGTTGTTGATT 65206

Qy   1069 TATCTAATAATTCCTCTTATCATAAATCAAGCCTATGAAATCGGGCTCTCCGVAAAATT 1128
Db  65205 TAAGCAAGATTCCTGTTCAAACAGCCTGCTCCAAGCAGAATCGGGCAGGCAGTATTC 65146
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2004, 15:49:34 ; Search time 151 Seconds
             (without alignments)
             7795.142 Million cell updates/sec

Title: US-10-698-235-1
Perfect score: 1656
Sequence: 1 aagtaatatccaccaatagsgg.........ggtgttkggaaraaataacg 1656

Scoring table: IDENTITY NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued Patents NA:*
                 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 5: /cgn2_6/ptodata/1/ina/pCTUS_COMB.seq:*
                 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	292.4	17.7	1830121	4	US-09-557-884-1	Sequence 1, Appli
2	292.4	17.7	1830121	4	US-09-643-990A-1	Sequence 1, Appli
3	292.4	17.7	1830121	4	US-10-329-960-1	Sequence 1, Appli
4	95.4	5.8	1335	4	US-09-328-352-2917	Sequence 2917, Ap
5	88.2	5.3	1377	4	US-09-252-991A-10413	Sequence 10413, A
6	88.2	5.3	1383	4	US-09-252-991A-10306	Sequence 10306, A
7	88.2	5.3	1410	4	US-09-252-991A-10749	Sequence 10749, A
c 8	74	4.5	1141	4	US-09-806-708B-22	Sequence 22, Appl
9	73.6	4.4	1299	4	US-09-543-681A-3189	Sequence 3189, Ap
10	73	4.4	10467	4	US-10-204-708-2	Sequence 2, Appli
11	71.6	4.3	19513	4	US-10-204-708-39	Sequence 39, Appl
12	70	4.2	5152	4	US-10-204-708-74	Sequence 74, Appl
c 13	69.4	4.2	19124	2	US-08-487-826B-13	Sequence 13, Appl
14	67.8	4.1	741	4	US-09-489-039A-1	Sequence 1, Appli
15	67.6	4.1	8961	4	US-10-204-708-80	Sequence 80, Appl
16	66.6	4.0	1055	4	US-09-806-708B-23	Sequence 23, Appl
17	66	4.0	6113	4	US-10-204-708-14	Sequence 14, Appl
18	65.8	4.0	6306	4	US-10-204-708-50	Sequence 50, Appl
19	64.6	3.9	658	3	US-08-998-416-595	Sequence 595, App
20	64.4	3.9	6866	4	US-10-204-708-20	Sequence 20, Appl
c 21	63.8	3.9	1141	4	US-09-806-708B-22	Sequence 22, Appl
22	63.6	3.8	5501	4	US-10-204-708-38	Sequence 38, Appl
23	63.6	3.8	6070	4	US-10-204-708-10	Sequence 10, Appl
24	63	3.8	9347	4	US-10-204-708-35	Sequence 35, Appl
25	62.2	3.8	6669	4	US-10-204-708-5	Sequence 5, Appli
26	62	3.7	6583	4	US-10-204-708-26	Sequence 26, Appl
27	62	3.7	11049	4	US-10-204-708-22	Sequence 22, Appl
28	61.8	3.7	636	3	US-08-998-416-1137	Sequence 1137, Ap
29	61.4	3.7	6669	4	US-10-204-708-6	Sequence 6, Appli
30	61	3.7	11049	4	US-10-204-708-21	Sequence 21, Appl
31	60.6	3.7	5455	4	US-10-204-708-33	Sequence 33, Appl
32	60.4	3.6	19124	2	US-08-487-826B-13	Sequence 13, Appl
33	60.2	3.6	11049	4	US-10-204-708-23	Sequence 23, Appl
34	60	3.6	5852	1	US-07-867-106-2	Sequence 2, Appli
35	59.8	3.6	6866	4	US-10-204-708-19	Sequence 19, Appl
36	59.4	3.6	3111	4	US-09-252-991A-10504	Sequence 10504, A
37	58.8	3.6	6156	4	US-10-204-708-60	Sequence 60, Appl
38	58.6	3.5	5666	4	US-10-204-708-30	Sequence 30, Appl
39	58.4	3.5	5562	4	US-10-204-708-63	Sequence 63, Appl
40	58.4	3.5	6040	4	US-10-204-708-69	Sequence 69, Appl
41	58	3.5	8093	4	US-10-204-708-32	Sequence 32, Appl
42	57.8	3.5	6317	4	US-10-204-708-11	Sequence 11, Appl
43	57.6	3.5	10619	4	US-10-204-708-3	Sequence 3, Appli
44	57.6	3.5	19513	4	US-10-204-708-40	Sequence 40, Appl
45	57.2	3.5	5610	4	US-10-204-708-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
;   the Haemophilus influenzae Rd Genome, Fragments
;   Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match   17.7%; Score 292.4; DB 4; Length 1830121;
Best Local Similarity   85.1%; Pred. No. 8e-54;
Matches 336; Conservative   1; Mismatches   57; Indels   1; Gaps   1;

Qy   8 ATCACAATAGSGGATCCACGAGCTTCTATTAGGTATCGTATTGGCTGCAGAGGATATCC 67

```
;          TYPE: nucleic acid
;          STRANDEDNESS: double
;          TOPOLOGY: linear
;     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match        17.7%;   Score 292.4;  DB 4;   Length 1830121;
Best Local Similarity  85.1%;   Pred. No. 8e-54;
Matches 336;   Conservative  1;   Mismatches  57;   Indels  1;   Gaps  1;

Qy          8 ATCACAATAGGGGATCCACGAGCTTCTATTAGGTATCGTATTGGCTGCAGAGGGATATCC 67
              ||| |||  :|||| |||||||||| |||||||||||||||||||||||||||||||||||
Db     941919 ATATCACAATGGGATCTGCGAGCTTCT-TTAGGTATCGTATTGGCTGCAGAGGGATATCCC 941977

Qy         68 AAAGGATTATCGCAAAGGCGATGAAATCAGCCGGATTGCCTAAAAGTGCGGTCAAAAACGA 127
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     941978 AAAAGATTATCGCAAAGGCGATGAAATCAGCCGGATTGCCTAAAAGTGCGGTCAAAAACGA 942037

Qy        128 GAAAGTTTCTTAGCGGGTGTCGCAGAACAAGAAGGCAAGCTAGTCACAAACGGCGGTCG 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942038 GAAAGTTTCTTAGCGGGTGTCGCAGAACAAGAAGGCAAGCTAGTCACAAACGGCGGTCG 942097

Qy        188 TGTACTTTGTGTGACTGCGTTAGGCGAAAGTGTATTTGAAGCACAACAACAAAAAGCGTTAAA 247
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942098 TGTACTTTGTGTGACTGCGTTAGGCGAAAGTGTATTTGAAGCACAACAACAAAAAGCGTTAAA 942157

Qy        248 ATTGGCTGAGCAAATTCAATGCTCTGGGCGTTTTATCGTCGAGACATTGGTTACAGGGC 307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     942158 ATTGGCTGAGCAAATTCAATGCTCTGGGCGTTTTATCGTCGAGACATTGGTTACAGGGC 942217

Qy        308 TGTGGAACGAGAACAAGCAAATAGTTAGAAATCTTGTTGAAATCCATATCATTATGATGAATTAATTCATTTAAGT 367
              |||||||||||||           |||||||            ||||||||||||||  ||||||||||||||||
Db     942218 TGTGGAACGAGAGTTGCAAAAATAATCCATATCATTATGATGAATTAATTCATTTAAGT 942277

Qy        368 ATTGTACAGGGTAGAATTGTATTTCCTAGGATTT 402
              ||||||||||| ||||  ||||     ||||
Db     942278 GATGAAAATTTAGGAATGAAATTTTAATTAAATTT 942312
```

RESULT 2
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
            Mark D. Adams
            Owen White
            Hamilton O. Smith
            J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
                      the Haemophilus influenzae Rd Genome, Fragments
                      Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville
;   STATE: MD
;   COUNTRY: USA
;   ZIP: 20850
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3 1/2 inch diskette
;   COMPUTER: Dell Pentium
;   OPERATING SYSTEM: MS DOS v6.22
;   SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/643,990A
;   FILING DATE: 23-Aug-2000
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/487,429
;   FILING DATE: 1995-06-07
;   APPLICATION NUMBER: 08/426,787
;   FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
;   NAME: Kenley K. Hoover
;   REGISTRATION NUMBER: 40,302
;   REFERENCE/DOCKET NUMBER: PB186P1C1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 301-610-5790
;   TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1830121 base pairs

RESULT 3
US-10-329-960-1
; Sequence 1, Application US/10329960
; Patent No. 6742927
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fragm
;   Patent No. 6742927
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,960
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin version 3.1
SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (4747)..(4747)
  OTHER INFORMATION: n equals a, t, g or c
FEATURE:
  NAME/KEY: misc_feature
  LOCATION: (9921)..(9921)
  OTHER INFORMATION: n equals a, t, g or c
FEATURE:
  NAME/KEY: misc_feature

LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)

OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c

```
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (145171)..(145171)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (145942)..(145942)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (147197)..(147197)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (150841)..(150841)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (152500)..(152500)
;   OTHER INFORMATION: n equals a, t, g or c
;   FEATURE:
;   NAME/KEY: misc_feature

Query Match         17.7%;    Score 292.4;    DB 4;    Length 1830121;
Best Local Similarity   85.1%;    Pred. No. 8e-54;
Matches   336;   Conservative    1;   Mismatches    57;   Indels    1;   Gaps    1;

Qy        8  ATCACAATAGSGGATCCACGAGCTTCATTAGTATCGTATTGGCTGCAGAGGGATATCC      67
              |||  : ||| |  |||||||||||||||||||||||||||||||||||||||||||||
Db   941919  ATATCACAATGGGATCTGCGAGCTTCT-TTAGGTATCGTATTGGCTGCAGAGGGATACCC   941977

Qy       68  AAAGGATTATCGCAAAGGCGATGAAATCAGCGGATTGCCTAAAAGTGCGGTCAAAAACGA  127
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   941978  AAAGGATTATCGCAAAGGCGATGAAATCAGCGGATTGCCTAAAAGTGCGGTCAAAAACGA   942037

Qy      128  GAAAGTTTTCTTAGCGCGGTGTCGCAGAACAAGCTAGTCACAAACGGCGGTCG         187
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942038  GAAAGTTTTCTTAGCGCGGTGTCGCAGAACAAGCTAGTCACAAACGGCGGTCG        942097

Qy      188  TGTACTTTGTGTGACTCGTTAGGCGAAAGTGTATTTGAAGCACAACAACAAAAAGCGTTAAA  247
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942098  TGTACTTTGTGTGACTCGTTAGGCGAAAGTGTATTTGAAGCACAACAACAAAAAAGCGTTAAA   942157

Qy      248  ATTGGCTGAGCAAATTCAATGGTCTCGGCGTTTTATCGTCGAGACATTGTTACAGGGC     307
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942158  ATTGGCTGAGCAAATTCAATGGTCTCGGCGTTTTATCGTCGAGACATTGTTACAGGGC    942217

Qy      308  TGTGGAACGAGAGCAAGCAAATAGTTAGAAATCTTGTTGAAATCTTTAATTAGTAGATAAAAAAT  367
              ||||||||||||||   || ||||| ||||| ||||| |||||||||| || |||||||
Db   942218  TGTGGAACGAGAGTTGCAAAAATAATCCATATCATTATGATGAAATTAATTTCATTTAAGT   942277

Qy      368  ATTGTACAGGGTAGAATTGTATTTCCTAGGATTT       402
              |||||||    |||||| ||||  |||||
Db   942278  GATGAAAATTTAGGAATGAAATTTTAATTAAATTT    942312

RESULT 4
US-09-328-352-2917
;  Sequence 2917, Application US/09328352
;  Patent No. 6562958
;  GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
;  SEQ ID NO 2917
;  LENGTH: 1335
;  TYPE: DNA
;  ORGANISM: Acinetobacter baumannii
US-09-328-352-2917
```

```
Query Match          5.8%;    Score 95.4;    DB 4;    Length 1335;
Best Local Similarity  57.6%;    Pred. No. 8.6e-12;
Matches   171;   Conservative    0;   Mismatches  126;   Indels    0;   Gaps    0;

Qy       35  ATTAGGTATCGTATTGGCTGCAGAGGGATATCCAAAGGATTATCGCAAAGGCGGATGAAAT     94
              ||||||||||||||||||||  |||| ||| |||||| ||| ||||||
Db     1038  AATCGGTATTGTACTTGCAGCAGAAGGTTACCCAGACACAGCGTGCGTAAAGGTGACGTTAT   1097

Qy       95  CAGCGGATTGCCTAAAAGTGCGGTCAAAAACGAGAAAGTTTTCTTAGCGCGTGTCGCAGA    154
              |||  | |||  |||| |||| ||
Db     1098  CTCAGGTATCGGTCAATCACCAGGAGGATACTAAAATCTTCCATGCAGCACTGCTACTCG   1157

Qy      155  ACAAGAAGGCAAGCTAGTCACAAACGGCGGTCGTGTGTACTTTGTGTGACTGCGTTAGGCGA   214
              || ||||||| ||||| | | || || | ||
Db     1158  TGAAGATGGGCATATCGTTACTTCTGCGGACGTGTACTTGCGTGACTGCTCTAGGTGA    1217

Qy      215  AAGTGTATTTGAAGCACAACAACAAAAAGCGTTAAAAATTGGCTGAGCAAATTCAATGGTCTGG   274
              |||||||||| |||||| ||| |||| |||  ||||||| |||||| |||| ||
Db     1218  TAGCGTTCTTGAAGCACAAATCAATGCTTTAGAAGTATGTGGCAAGTAACCTTCACAGG    1277

Qy      275  GCGTTTTTATCGTCGAGACATTGGTTACAGGGCTGTGGAACGAGAACAAGCAAATA    331
              | ||||| |||||| |||||||||  | ||
Db     1278  TATGCAATACCGCAGTGACATTGGTTACCGTGCTATTGCTGCTGCGTGAAAAAGCTGAATA   1334

RESULT 5
US-09-252-991A-10413
;  Sequence 10413, Application US/09252991A
;  Patent No. 6551795
;  GENERAL INFORMATION:
;  APPLICANT: Marc J. Rubenfield et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE: 1999-02-18
;  PRIOR APPLICATION NUMBER: US 60/074,788
;  PRIOR FILING DATE: 1998-02-18
;  PRIOR APPLICATION NUMBER: US 60/094,190
;  PRIOR FILING DATE: 1998-07-27
;  NUMBER OF SEQ ID NOS: 33142
;  SEQ ID NO 10413
;  LENGTH: 1377
;  TYPE: DNA
;  ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10413

Query Match          5.3%;    Score 88.2;    DB 4;    Length 1377;
Best Local Similarity  56.8%;    Pred. No. 3.2e-10;
Matches   162;   Conservative    0;   Mismatches  123;   Indels    0;   Gaps    0;

Qy       37  TAGGTATCGTATTGGCTGCAGAGGGATATCCAAAAGGATTATCGCAAAGGCGATGAAATCA     96
              ||||||||||||||  ||||| ||||||  ||||| ||| |||||| |||
Db     1079  TGGGCGTGGTACTGGCCGCCGGGGGCTATCCGGGCGACTACCGCGGGGCCATCGCAGA   1138

Qy       97  GCGGATTGCCTAAAAGTGCGGTCAAAAACGAGAAAGTTTTCTTAGCGCGTGTCGCGAGAAC    156
              | || |||  || ||||| ||   || | ||| || | ||  |||
Db     1139  AAGGCCTGGCCGCAGGCCGCCGCGCTGGACGGACGGCAAGGTGTTCCACGGCCGGCTGA   1198

Qy      157  AAGAAGGCCAGCTAGTCACAACAAAAACGGCGGTCGTGTACTTTGTGTGACTGGCGTTAGGCGAAA   216
              || ||||| |||| | || || || || | ||
Db     1199  AGGATGGCCAGGAGCCCACCTCCGGCGTGCTCTGTGCCACCACCGCCGAGA    1258

Qy      217  GTGTATTTGAAGCACAACAACAAAAAGCGTTAAAATTGGCTGAGCAAATTCAATGGTCTGGGC   276
              |||||||||| |||||| |||| ||  |||| |||  ||||||
Db     1259  GCGTGTCCGGCGCCGGCCCAGCAACAGCGTCCTATCGCCTGGGCCGGAGAGATCCGCTGGAACGGCT   1318

Qy      277  GTTTTTATCGTCGAGACATTGGTTACAGGGCTGTGGAACGAGAAC    321
              |  |||| ||| |||||| |||  | ||
Db     1319  GCTTCTACCGGCAAGGACATCGGCTACCGTGCCATCGCCCGCGAGC    1363
```

RESULT 6
US-09-252-991A-10306
; Sequence 10306, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10306
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10306

Query Match   5.3%;   Score 88.2; DB 4; Length 1383;
Best Local Similarity   56.8%;   Pred. No. 3.2e-10;
Matches 162; Conservative   0; Mismatches 123; Indels   0; Gaps   0;

Qy   37 TAGGTATCGTATTGGCTGCTGCGAGAGGATATCCAAAGGATTATCGCAAAGGCGATGAAATCA 96
            ||| |||| || ||||||| ||||||| | |||| ||| | ||| || ||||| |||
Db   957 TGGGCGTGGTACTGGCCGGAGGCGCCGGCTATCCGGCGACTACGCCAAGGGCGAGGTCATCG 1016

Qy   97 GCGGATTGCCTAAAAGTGCGGTCAAAAACGAGAAAGTTTTCTTAGCGGGTGTCGCAGAAC 156
            |||| |||| ||| || ||||| || ||| |||| | |||||||||||||||| ||
Db   1017 AAGGCCTGGCCGAGGCCGCCGCGCGCTGGACGGCAAGGTGTTCCACCGCGGCGAGGTCATCG 1076

Qy   157 AAGAGGCAAGCTAGTCACAAACGCGGTCGTGTACTTTGTGTGACTGCGTTAGGCGAAA 216
            || || || ||| | || |||| || ||| ||| ||||| ||||||||||| ||||
Db   1077 AGGATGGCCAGGAGTCACCTCCGGCGGCCGCCATCGGCGAGA 1136

Qy   217 GTGTATTTGAAGCACAACAAACAAAAGCGTTAAAAATTGGCTGAGCAAATTCAATGGCTGGGC 276
            ||||| | |||| || |||| ||||||||| | |||||||| ||||| |||||||
Db   1137 GCGTGTCCGCCGCCCAGCAACAGCGCCATCGCCTCGGCCCGAGAAGATCCGCCTGGAACGGCT 1196

Qy   277 GTTTTTATCGTCGAGACATTGGTTACAGGGCTGTGTGAACGAGAAC 321
            ||||| ||| |||||||| |||||||||||| |||||||| ||
Db   1197 GCTTCTACCGCAAGGACATCGGCTACCGTGCCATCGCCCGCGAGC 1241

RESULT 7
US-09-252-991A-10749/c
; Sequence 10749, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 10749
; SEQ ID NO 10749
; LENGTH: 1410
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10749

Query Match   5.3%;   Score 88.2; DB 4; Length 1410;
Best Local Similarity   56.8%;   Pred. No. 3.2e-10;
Matches 162; Conservative   0; Mismatches 123; Indels   0; Gaps   0;

Qy   37 TAGGTATCGTATTGGCTGCTGCGAGAGGATATCCAAAGGATTATCGCAAAGGCGATGAAATCA 96
            ||| |||| || ||||||| ||||||| | |||| ||| | ||| || ||||| |||
Db   365 TGGGCGTGGTACTGGCCGGAGGCGCCGGCTATCCGGCGACTACGCCAAGGGCGAGGTCATCG 306

Qy   97 GCGGATTGCCTAAAAGTGCGGTCAAAAACGAGAAAGTTTTCTTAGCGGGTGTCGCAGAAC 156
            |||| |||| ||| || ||||| || ||| |||| | |||||||||||||||| ||
Db   305 AAGGCCTGGCCGAGGCCGCCGCGCGCTGGACGGCAAGGTGTTCCACCGCGGCGAGGTCATCG 245

Qy   157 AAGAGGCAAGCTAGTCACAAACGCGGTCGTGTACTTTGTGTGACTGCGTTAGGCGAAA 216
            || || || ||| | || |||| || ||| ||| ||||| ||||||||||| ||||
Db   245 AGGATGGCCAGGAGTCACCTCCGGCGGCCGCCATCGGCGAGA 186

Qy   217 GTGTATTTGAAGCACAACAAACAAAAGCGTTAAAAATTGGCTGAGCAAATTCAATGGCTGGGC 276
            ||||| | |||| || |||| ||||||||| | |||||||| ||||| |||||||
Db   185 GCGTGTCCGCCGCCCAGCAACAGCGCCATCGCCTCGGCCCGAGAAGATCCGCCTGGAACGGCT 126

Qy   277 GTTTTTATCGTCGAGACATTGGTTACAGGGCTGTGTGAACGAGAAC 321
            ||||| ||| |||||||| |||||||||||| |||||||| ||
Db   125 GCTTCTACCGCAAGGACATCGGCTACCGTGCCATCGCCCGCGAGC 81

RESULT 8
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match   4.5%;   Score 74; DB 4; Length 1141;
Best Local Similarity   11.0%;   Pred. No. 3.7e-07;
Matches 110; Conservative 400; Mismatches 474; Indels 13; Gaps   4;

Qy   358 GATAAAAAATATTGTACAGGTGAGAATTGTATTTCCTAGGATTTAGGATTTTGTTAGGG 417
            :::: :|:: : : :   : : :: || :: : :::   : ::
Db   56 SRKWTWARMYCKYRRWYNNKSRWWKGWYKKKWYBCANNTSBRYHARRWKDMKTAYBMTMT 115

Qy   418 CAACGTTTACGATTGCTCTGACAATAAATTAGAATTATTATTTTTGTTTACTTTATCGAGGT 477
            : :||:::: :: ::: ::   : : |:: :   ::::::: :   :::: :: :
Db   116 NKWGKTGWRHRYWRWRAMBDTVDHHYVTAMNNAWTTMCMDKDDKRTRWWWKKNNNATGW 175

Qy   478 TATATCAACTTATGCGACAATTTGTCATCGTAGTAGTATTATATTCTGCAATTCTTCTTCAT 537
            ::: :::: :::: :   :: :   ::: :: : : :::: :: ::: :::::::
Db   176 DDDTKYHMWNNNGCBTVTWMVRYKTDRDWSBKRMNYGMBWWKNWSYDVTYYWWVWDDMCK 235

Qy   538 TAGAAGTTATTTATAGAAAATTATTTAATATTCTAGTATTGAGAGATACACTGAAAGTT 597
            ::: ::::   : ::   : ::::   ::   :: :: :: :::
Db   236 RKVRRWVRTRGRMRNYMVAWBTAHRRRYNNGWTBAMAYRRWTMNNNNNNAKAMCKRAKYW 295

Qy   598 ATTTGTCGTTGTTGTTGTTGTTGTTGTTGTTGTTTTCAAAATAGAATTACAGAA 657
            :: : :: :::: ::: ||::   : : ::: ||:::::::
Db   296 GWNRABVNSTCTTWKSKTTWKRRTSCWANNCRAGDANKDHKWKWKSAAMGVWNNNNNNW 355

Qy   658 TATTAGTTGGCGCTTTATTTGCAATAAGTATTGTTGTTAATAAATGTACATTATGC-AGTA 716
            ::::|::| :::::   : :::   :: :   :: :::::: :   :: ::
Db   356 TYKKARHBARWDWVHSAWKKWHANAAHYSRKKWTBYKRKTMVNNNNGTTMWKRMWAWYW 415

```
Qy   717 TACCAATCTTGGATTGGACCTGTTAATTACTCACTTGCATTTAAAGAAATTAATGAGATA 776
             :  ::::::|:  :|  |:   ::: |:   ::|:::||:    ::::|:|
Db   416 KMDMDWBGTYNNNNNGGRTYYGWTKNKKMWTYYKWKANNCKWRAWDHKTCTHNNTTWWKM 475

Qy   777 ACAAATGCTGCTTAACAATGATAGATAAATTCATATATCCATTGTTATTTGGTTTATTT 836
             :  ::::|:   :   ::|   |   :|:   :::
Db   476 KTYWNNCYWKSMTNGKSHRBAAAVYTWYMWWWRYA---HANNNNWDYWWKACTWYKYB 531

Qy   837 GAAGTTGCTGCTGTGTTTTAAGTTAAGCTTCATAAAAGAGAAAGTATATAAACTTCTTGG 896
             :::   :|   :|:::|   :::||:   :|::: :::
Db   532 VCSKWNNYAAWYTKSSWNYTSRYRWKTNNSWRWRSDTRSMGRANNYARABHYGYKWNT 591

Qy   897 ATTTTTGACTTTATTTTTATCGTGTGATGATGTATGTTTTTGTTCGAGCGTATACAACA 956
             :  ||:::   :||:::   : :::|   :||  |||
Db   592 RWWBW---SHTWBHBRAGAAHYWMBMMYBAKCHCMKAWYKAKKYAGAGGSNNNNNNNN 647

Qy   957 AAATCCCATGAGCGTTTTATTTTCACCTAAACACTGTTTTATTCTCGATTAAAATCCAATTAT 1016
             :  :  ||: ::  : :  :::  :   :: :::||
Db   648 NNNNNNNATCARDDYYAASRWYAMANAKWYYYKBAANNAYYTHANNWWGCWNNATDTRRRT 707

Qy  1017 TTATCGTTGGGTTATTTTATAGGACGAATTGTTCCTTATGAGATATTTTTCTTTATCTAAT 1076
             ||:   ::|   :   :: :::   :||||:   ::: ::: |::||||  :::
Db   708 MWKNNNNNAGTWKNNNNNNAKNASAAKNYAAAAVKAAKKHWRWANKWAMRGWHADAAAB 767

Qy  1077 ATTCCTCTTTATCATAAATCTAAGCCTATGAAATCGGGCTCTCCGVAAAATTMARGAATA 1136
             |:   :||||:   : :   :||:   :|
Db   768 TTDKRRNGAYTKYTTNNNNTYRGVVTNTAARDGWANNNNNNNNNNNGWSDMWVTWW 827

Qy  1137 TAATTTTAATTAATAGGGGAAAGTGCGACCTCAAGTCATTTTAGTGCTTT----TGGTTA 1192
             :::::::   :::|||   :::|   :|:::: ::|   ::|:||
Db   828 AYANYGTNNNNNNNNNAYAWWTNKWYYTTDDRWRBAYTNNNNNNRMAYYGAYADDYAYY 887

Qy  1193 CGGGAGAAAAACATCTCCTTTTTAGATAGCTTAAAATAAAAATCAGGAGCTCTTGTTGG 1252
             :::   |||:|   :  ::|:   :|||:   :   :::   :|||:   : :::
Db   888 MSDTCDAWMKWDATKMNNATTYNRGTAWRTNNNNNMTKTKYYBHAAWNNNNNNNGKMCT 947

Qy  1253 TAAAACTTATTCAGGAGGAAAGCTAACAGCAATTCTTACCCAATGTTTTTAATCGCAAT 1312
             : :   :|:||   : :   |:   ::|||   :   ::|:||   ||:::
Db   948 AHTWWVCKATKTTKGCWMNCTTTCRKYKNNCTWYTWMTTTRTTWYAATRWKTNNATGSMT 1007

Qy  1313 YCCTTAYCCAAATGGAATACAACAGATAGCTAAAGGA 1349
             |::: ::||   :::|::   ||::|||:
Db  1008 RCNATGWKNNNNYWTGWKTRWTAYRMATRWMKAlWWKVM 1044
```

RESULT 9
US-09-543-681A-3189
; Sequence 3189, Application US/09543681A
; Patent No. 665709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABII
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3189
; LENGTH: 1299
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-3189

```
Query Match       4.4%;    Score 73.6;   DB 4;   Length 1299;
Best Local Similarity   55.5%;   Pred. No. 4.7e-07;
Matches   181;   Conservative   1;   Mismatches 140;   Indels   4;   Gaps   2;

Qy     1 AAGTAATATCACAATAGSGGATCCACGAGCTTCATTAGGTATCGTATTGGCTGCGAGAGG 60
             ||:   ||||   ::|| |||||   : ::::|   || ::||
Db   969 AGGTAAAGACTCCATTTGGGATCCCGGCCCGCCCCT-TTAGGTGTTGTGTAAATGGCCGCCGGTG 1027
```

```
Qy    61 GATATCCAAAGGATTATCGCAAAGGCGATGAAATCAGGCGATTGCCTAAAAGTGCGGTCA 120
             | ||||:|   |   ::|:   :||||||   || ::| || |||||::||   |||
Db  1028 GCTATCCTGATAACTACCGCCCAAAATGTATATTATTGGGGACTAACAGCAACATCTTCTT 1087

Qy   121 AAAACGAGAAAGTTTCTTAGCGGGTGTCGCAGAACAG---AAGGCAAGCTAGTCACAA 177
             |||   ::||   :  ||:   :   :   ::::|   ::|   |||   :::   ::
Db  1088 CAACAGCTAAGGTATTCAAGCAGGAACAAAATTAACGGCTCAAGGTGAAGTTGTTACAG 1147

Qy   178 ACGGCGCTGTGTACTTGTGTGACTGCGTTAGGCGAAAGTGTATTGAAGCACAACAA 237
             |   ::|   |||   : :|| |:::||   | :: |   |||   |||:|
Db  1148 CAGGAGCCGAGTATTATGTGCAACTGCTTAGGGAAGAATATCGAACAAGCTCAAAAGA 1207

Qy   238 AAGCGTTAAAATTGGCTGAGCAAATTCAATGGTCTCGGCGTTTTTATCGTCGAGACATTG 297
             |||   ::|   ||:   : |   |||   :|||||||   :||||
Db  1208 ATGCCTATGCTTAGCCTTTAGCCAAAAGTATTCACTGTATGCTGTTTCTATCGCCATGATATCG 1267

Qy   298 GTTACAGGGCTGTGGACAGACAACAA 323
             |||   :|   |||   ||::|
Db  1268 GCTATCGTGCCATTGCGCGTTTAAAA 1293
```

RESULT 10
US-10-204-708-2
; Sequence 2, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 2
; LENGTH: 10467
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-2

```
Query Match       4.4%;    Score 73;   DB 4;   Length 10467;
Best Local Similarity   44.3%;   Pred. No. 1.1e-06;
Matches  379;   Conservative   1;   Mismatches 471;   Indels  4;   Gaps  2;

Qy   337 AAATCTTGTTGAATTAATTAGATAAAAATATTGTACAGGGTAGAATTGTATTTCCTA 396
             ||::   ||||| ||||   ||::||
Db  1883 AAAATTTTTTTTTTTTTTGAATTTTATTTGGGTTTCGTTGTGTGTGGTTAATTGTGTTAA 1942

Qy   397 GGATTAGGATTTGTTAGGGCAACGTTACCAGTTCTGCGACAATAAATTAGAATTATT 456
             ||||||   |: ||||   : ||   :||::   ||  || ::|||
Db  1943 TAAATATGTTTTAGTTTATTGTGGATTTGATTTTAAAAATTTTTTTTAG 2002

Qy   457 ATTTTTGTTGACTTATGCAGGTTGATATTATCATCGTCTAGTAGTATTAT 516
             |||::|   :|   ||::||||
Db  2003 GGTTTGGTTGATATAGTAAAGGTTTTTTTATTTTTTTTTGTTTTTTTTTAT 2062

Qy   517 ATTCTGCAATCTTCTTTCATTAGGTATCGTATTGGCTGCGAGAGG 576
             |||:   || ||:|||
Db  2063 TTTTATAATTTTTTATAGTT--ATTGTATTAACGGATTATTTTTTTATATAATTAGTT 2120
```

```
;    FEATURE:
;    OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-39

Query Match                4.3%;     Score 71.6;  DB 4;  Length 19513;
Best Local Similarity   47.4%;     Pred. No. 2.6e-06;
Matches 287;  Conservative  0;  Mismatches 309;  Indels 10;  Gaps  2;

Qy     337 AAATCTTGTTGAATTAATTAGATAAAAAAATATTGTACAGGTAGAGATTGTATTTCCTA 396
           |||||  || |||||||    |   |   ||||| |||||| |||||  ||||||
Db    8938 AAATATTATTGAATTAATATTTAAATGGGTAGAAGTTGTATGTGAATTATAGT 8997

Qy     397 GGATTTAGGATTTTGTTAGGGCAACGTTACGATTGCTCTGACAATAAAATAGAATTATT 456
           | ||||||| |||||||||||   |   |||||||  ||||||||||||||||
Db    8998 TTAATAAAGTAGTTAGTAAATGAAGTTTTATTTGTTTTTAGGAAATTAGTTTTTTA 9057

Qy     457 ATTTTTGTTACTTTATGAGTTATATCAACTTATGCGACAATTGTCATCGTAGTATTAT 516
           ||||||||   |||||||| |||| |||  | ||||    |||||||||||||||||
Db    9058 GTTTTTTTTATAGAATGATGTATTAAAAGTAAAATGTATTATATATGTTTGTATTAT 9117

Qy     517 ATTCTGCAA----TTCTTCTTCATTAGAAGTTATTATAGAAAATTATTTAATATTTCT 572
           ||||||||    |||| |||||||||||| ||||| |||  |||||||||||||
Db    9118 AATAAAAAGACGTTTATATATTTTTGTATATAAATAAAAAAGACGTTTATATATTTTT 9177

Qy     573 AGTATTGAGAGATACACTGAAAGTATTGTCTCGTTTGTTGTTTTGTTTGTTTTGTTGTT 632
           |||  ||||| ||||||||    |||||||| ||||| ||| ||||||||
Db    9178 TGTATTAAAAGTAAAAAGATTTTATATATTTTTTGTTTTTGTTTATTTGTTATTA 9237

Qy     633 TTTTCAAAATATAGAATTACAAGAGATATTAGTTGGCGCTTATTGCAATAAGTATTGTT 692
           |||||      |||| ||||| |||  |||||  |||||  ||||| |||||||||||
Db    9238 TTTTTATTTTATATATTTTTTTAATATTTTGAAAAGTTTAAATTTGTAGAGAAGGT 9297

Qy     693 GTTAATAATGTACATTATGCAGTACCAATCTTCGATTGTTTTTTGTTTTAGGTTGTTGTATG 752
           ||||||||  |||   |||   |||   |||   || ||||| ||||
Db    9298 ATAAATTGAGTATAGTAATAGTTTATTGGTTTTTGTTTTAGGTTGTTGTATG 9357

Qy     753 GCAATTAAAGAAATTAATGAGAATAACAAATGCTCGCTTAACAATGATAGATAAAATCATA 812
           ||||                   |   |||||||  |   ||||||
Db    9358 ATATTTTATTTGTTTTATGTTTTATTAAGTAAATATATATTTGT-----ATTTTTT 9411

Qy     813 TATCCATTGTTATTGTATTTGGTTTATTGTTGTGTGTTTTTAAGTTTTCATAAAA 872
           || ||| ||||| ||  ||  ||||| ||||  |   |||  |||||
Db    9412 GTTTTTTTTTGGTTGAATTATTTGAAAGTTGTGTGTTTTATTTTTATTTTTGAAA 9471

Qy     873 AGAAAGTATATAAACTTTCTTGCGATTTTTTTTTTTTTTTTTTTATGCTGTGATGATGTAT 932
           || ||||| ||||| ||||| ||||| |||||||||||||||||| |||||
Db    9472 ATTTTAGTAGTTAATTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTTTGTTTT 9531

Qy     933 GTTTTT 938
           ||| ||
Db    9532 GTTATT 9537
```

RESULT 11
US-10-204-708-39
; Sequence 39, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 39
; LENGTH: 19513
; TYPE: DNA
; ORGANISM: Artificial Sequence

```
Qy     577 TTGAGAGATACACTGAAAGTATTGTCGTTTGTGTTTGTTTGTTTGTTTGTTTTT 636
           |||||               ||||||   |||
Db    2121 TTTAAGGTTTTAAGGATTTTTTTTTTTTTATAATTATTTTTAGGTTGTGTTTAATTAT 2180

Qy     637 CAAAATA--TAGAATTACAAGAATATATTAGTTGGGCGCTTTATTGCAATAAGTATTGTGT 694
           |||||   |||||||  |||| |||| |||||||
Db    2181 TGTTTTAATTTAAATTAGTTTTTGAAGTAAGAGTATTTAAATTTATATGTTTAGTTAT 2240

Qy     695 TAATAAATGTACATTATGCAGTATACCAATCTTGGATTGGACCCTGTTAATTACTCACTTGC 754
           ||||||  |||| |||||  ||||  |||   | ||| |||
Db    2241 TATTTTGTGTGGGAGTTTTAAGATTTGAATTTTTAATTATGTCGATAGTTGTTTGTTATTTA 2300

Qy     755 ATTTAAAGAAATTAATGAGATAACAAAATGCTGGCTTAACAATGATAGAATAAATTCATATA 814
           ||||| |||||||||| ||||     ||||  ||| ||||||||||||
Db    2301 ATATTTTGTATATTAATTAATTAATAGTAAAATTTGAATTAAATAATTAGTTTTTTAAAATT 2360

Qy     815 TCCATTGTTATTGGTTTATTGGTTGAAGTTGCTGTCGTGTTTTTATTTAAGTTTAAGTTTCATAAAAAG 874
           |||||  |||||||   |||||||  |||||||||||||||||||| ||||||||| |||
Db    2361 TAGTTTTTTTTTTTTGTAGTTTTAATGGTATTATTTTTTTTTTTTTTTAGTTTTATTAAG 2420

Qy     875 AAAAGTATATAAACTTTCTGGATTTTTGTTTTTGTTTTATTTATGCTGTGATGATGTATGT 934
           |||||||| |||  |||| || |||| ||||||
Db    2421 GTTAGAAATTTGCGTTTATTTTGTTCGTTTTTTGTTTTTTGGATTTTATTA 2480

Qy     935 TTTTGTTGCAGCGTATACACAACAAAATCCCATGAGAGCGTTTTATTCTGTTGTTTGTATTGTGTTTAAAATGTGTTTTTT 994
           | |||||  |||||   ||||||||   |||| ||||||  ||||   ||||| ||| |
Db    2481 GGTAGTTGTTAAGTTTGTAAGTTTTTTTTTTGTATTGTGTTTTTTTTGTGTTTTAAAATGTGTTTTTT 2540

Qy     995 TTCTCGATTAAAATCCAATTATTATTCGTTGGGTTTATTTATAGGACGAATTGTTCCTTA 1054
           |||| ||||||     |||||| || ||||||||
Db    2541 TTTAGTTTTTAGTTTTTGTTTTTGTTTTATTGGGTTTATTTTTGTAAAGGTTTATAATTTTTTTA 2600

Qy    1055 TGAGATATTTTCTTTATCCTCTTATCATAAAATCTAAGCCTATGAAAATCGGG 1114
           ||||| |||     |  |||     ||||   || ||||||
Db    2601 TTTTTTTATATGTTAATTTAGGTGGTATATTTTTTTTTATAGTTAGTTATTATTTTTG 2660

Qy    1115 CTCCCCGGVAAAAATTMARGAAATAATAATATATTTAATTAATGGGGGAAAGTGCGACCTCCAAGTCA 1174
           || ||||||        ||   | |  |||||  ||  |||||
Db    2661 ATTTTTTTTATAGAGTTTTTTTTTAAAATGTTATTTGTTTTTTGTTTTTAGGAGAGTTTTTTGATTTTTATT 2720

Qy    1175 TTTTAGTGCTTTTGG 1189
           |||| |||||||||
Db    2721 TTTTAGTTTTTTGG 2735
```

RESULT 12
US-10-204-708-74
; Sequence 74, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 74
; LENGTH: 5152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-74

Query Match   4.2%; Score 70; DB 4; Length 5152;
Best Local Similarity   47.0%; Pred. No. 4.1e-06;
Matches 283; Conservative 0; Mismatches 315; Indels 4; Gaps 2;

Qy   380 AGAATTGTATTTTCCTAGGATTAGGATTTTGTTAGGGCAACGTTACGATTGCTCTGAC 439
              ||| |||||||||||||||||||| ||||||||||||||||||
Db   176 AGATTAGTTTAGTATAAGATTTTAAAATTTTTTAAATTTTTTTTTTTTAGGGAAAGT 235

Qy   440 AATAAAATAGAATTATTATTTTGTTACTTTATGAGGTTATATCAACTTATGCGACAAT 499
              |||||||||||||||||||||||||| |||||||||
Db   236 GGAAAAATAAGGAACGGATTTTATTTTTTTTTTTTTAGAGGAATAAATTTAGAATAAGAT 295

Qy   500 TGTCATCGTAGTATTATATTCTG-CAATTCTTCTTCATTAGAAGTTATTATAGAAAAT 558
              |||||||||||||||||||| |||||||||||||||||
Db   296 ATTTAAATGTAGGATTTGAGAATGTTGATGTTTTATTTTAAAGTAGGTGTTTTATATTAT 355

Qy   559 TATTTAATATTTCTAGTATTGAGAGATACACTGAAAAGTTATTTGTCGTGTTTGTTTTTTG 618
              |||||||||||||||||||||||||||||||||||||
Db   356 ATATTTTAGAAATTTTTATTTATAAAATTTATTTTATGTTTTTTTATTTTTTATAT 415

Qy   619 TTTGTTTGTTTGTTTTTTCAAAATATAGAATTACAAGAATATAATGTTTTTATTTTTTAAGTTG 678
              |||||||||||||||||||||||
Db   416 ATTTTTATTTGTTTAGTGTTTTTTACGAAAAAATAAATGTTTTTTATTTTTTTAAGTG 475

Qy   679 CAATAAGTAATTGTTGTTAATAAATGTACATTAATGTCAGTATACCAAATCTTGGATTGGACCTG 738
              |||||||||||||||| ||| ||||||||||
Db   476 TTATAATTTTTTAGTTAATATTTCGTAAGAAATTTTTAGAGATTTTTATTAGGTTAGATTTT 535

Qy   739 TTAATTACTCACTTGCATTTAAAGAAATTAAAGATGAGATAACAAATGCTGGCTTAACAATGA 798
              |||||||||||| ||||||||||||||||||||||||||
Db   536 TTATTTTTTAAAATAGGTGTTTTTAAAATATTTTTTGAAAGGTTTTAAGTTTTTTTATT 595

Qy   799 TAGATAAATTCATATATCCATTTGTTATTTGGTTTTATTTGAAGTTGCTGTGTTTTTAAGTT 858
              T---TTTATTGTAATGTTTTTTTTTGAAGGTTTTTGAGATAGGTGGGATGGTTTTGTTTGTGT 652
Db   596

Qy   859 TAAGTTTCATAAAAGAAAAGTATATAAAGTTTTTTTGGATTTTCTTCTTTGGATTTTTATTTTTATG 918
              |||||||||||||||||||||||||||| |||||||||||||||||||
Db   653 TAAGGATTGTTATATGTTGTTTTTTTGGGTTTTTTTTTTTTTTTTTAATTTTGTTTTTATT 712

Qy   919 CTGTGATGATGTATGTTTTTGTTCGAGCGTATACAACAAAATCCCATGAGCCGTTTTATTT 978
              |||| ||||||||||||||||||||||
Db   713 GAAAAATGGATTTGGGATAGAATTTAAAGTATTTTTTAAGTTGGAGTTAGAGTTTTATTTT 772

Qy   979 CA 980
              ||
Db   773 TA 774

RESULT 13
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 5593827
; GENERAL INFORMATION:
;   APPLICANT: Sim, Kim L.
;   APPLICANT: Chitnis, Chetan
;   APPLICANT: Miller, Louis H.
;   APPLICANT: Peterson, David S.
;   APPLICANT: Su, Xin-zhaun
;   APPLICANT: Wellems, Thomas E.
;   TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Knobbe Martens Olson & Bear
;   STREET: 620 Newport Center Drive 16th Floor
;   CITY: Newport Beach
;   STATE: California
;   COUNTRY: US
;   ZIP: 92660
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/487,826B
;   FILING DATE: 10-SEP-1993
;   CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;   NAME: Israelsen, Ned
;   REGISTRATION NUMBER: 29,655
;   REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (619) 235-8550
;   TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 19124 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-487-826B-13

Query Match   4.2%; Score 69.4; DB 2; Length 19124;
Best Local Similarity   44.1%; Pred. No. 7.8e-06;
Matches 289; Conservative 0; Mismatches 366; Indels 0; Gaps 0;

Qy   343 TGTTGAATTAATTAGATAAAAATATTGTACAGGGTAGAATTGTATTTTCCTAGGATTT 402
              |||||||||||| |||||||||||||||||
Db 15985 TGTATATGTGCATGTAATTAATTTAATTTTATTTTTTTAATTAAATTAATTTTTTATT 15926

Qy   403 AGGATTTTGTTAGGGCAACGTTACGATTGCTCTGCACAAATAAAATGAGAATTATTATTTTT 462
              ||||||||||||||||| |||||||||||||
Db 15925 TTTTTTTATTTCATTAAATTTTTTTTAATTTTTTTTTTTTATTTTTTAATTAAATTTTTTATTTAT 15866

Qy   463 GTTACTTTATGAGGTTATATCAACTTATGCGACAATTGTCATCGTAGTATTATATTCTG 522
              ||||||||||||||||||||||| |||||||||||
Db 15865 TATTTTTTTTTATTTAAATTAAATTTTTTTTATTTAATTATTCGTAGTATTATTAAATAAATTT 15806

Qy   523 CAATTCTTCTTCATTAGAAGTTATTATAGAAAAATTATTTAATATTTCTAGTATTGAGA 582
              ||||||||||||||||||||||||||||||||
Db 15805 TTTTTATTTATGTATATTTTTTCTTTTCTTTTGTTTTGTTTTTTGTTTTTATGATATATATGA 15746

Qy   583 GATACACTGAAAGTTATTTGTCGTGTTTGTTGTTTTTATTTGGCGCTTATTTGTTTTTTCAAAAT 642
              || ||||||||||||||||||||
Db 15745 TATATATTTTTTAAATGTTTTTTAAATATTTTTTTCTTTTTTTTTGTTTTTATATACATTTTT 15686

Qy   643 ATAGAATTACAAGAATATAAGTTGGCGCTTATTTGCAATAAGTATTGTTGTTAATAATG 702
              ||| ||||||||||||||||||||||||||
Db 15685 TTTTTTTTTATAAGTTTTTTTTTTCTCTTTTGTTTTTTTTTATTTTTTTATAAATCATTTTT 15626

Qy   703 TACATTATGCAGTATACCAATCTGGATTGGACCTGTTAATTACTCACTTGCATTTAAAG 762
              |||||||||||||||| ||||||||||
Db 15625 TTTTATATAAAATTTTTTTTAATTTTTTTAATTTTTTTTTTTCATTTTTAATTCTATC 15566

Qy   763 AAATTAATGAAGATAACAAATGCTGCTTAACAATGATAGATAAATTCATATATCCATTGT 822
              ||| |||||||||||||||||||||
Db 15565 AAAATTTTATATTTATTATTTTTTAATTTTTTTAAATTTTTTCTCCCTTTTTTTTT 15506

Qy   823 TATTTGGTTTTATTTGAAGTTGCTGTGTTTTTAAGTTTTTTTCATAAAAGAAAAGTAT 882

```
Db   15505  TTTTTTTTTTTTTATTTTTATTTAATAAATTTTTTTTTTCATTTTTCTTTTTTTCATTTTA  15446
              | ||||  ||||||            ||||||                  |||||||
Qy    883  ATAAACTTCTTCTTGGATTTTTGACTTTATTTTTTTATGCTGTGATGATGTATGTTTTTGTTC  942

Db   15445  ATAAATTTGTTTTTATATTTCTTTTTAATAAATAACATATATATAAAATATATATATATA  15386
              ||||| |||||||              ||||||
Qy    943  GAGCGTATACAACAAAATCCCATGAGCGTTTATTCACCTAAACACTGTTTATTC  997
              |||     ||||  |  |||||         ||||||    ||||||   |||||||
Db   15385  AAACACATACAAGTACATCCATATATATATCCTTAACCCACAAATACATATAC  15331
```

RESULT 14
US-09-489-039A-1
; Sequence 1, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1
; LENGTH: 741
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1

  Query Match   4.1%; Score 67.8; DB 4; Length 741;
  Best Local Similarity 54.2%; Pred. No. 7.4e-06;
  Matches 160; Conservative   0; Mismatches 132; Indels 3; Gaps 1;

```
Qy    37  TAGTATCGTATTGGCTGCGAGAGGGATATCCAAAGGATAATCGCAAAGGCGATGAAATCA  96
            |||  |||     |||||  |||| ||||  |   |||||  ||||
Db   440  TGGCGTGGTGGTTGCCGCGCCGGTGCGCTATCCGGGCAGCTACAATACGGGCGATAGAGATCT  499

Qy    97  GCGGATTGCCTAAAAGTGCGGTCAAAAACGAGAAAGTTTTCTTAGCGGGTGTGTCCAGAAC  156
            |  ||||||||||||||||    ||  |   | |||     ||   || || ||||
Db   500  ACGGCCTGCCGCAGCAGAAGTCGAGACCGGTAAGGTCTTCCCACGCGCACCAAACTGT  559

Qy   157  AAGAAGGCAAGC---TAGTCACAAACGGGGTCGTCGTATCTTGTGTGACTGGCGTTAGGCG  213
            ||||||  |||          |||  ||| ||| ||  |||  || ||||||||
Db   560  CTGACGACCAGCCGCGTGGTCACCAACGGGCGGGCCGTGCGTGGCCGCCGCGGGGTG  619

Qy   214  AAAGTGTATTTGAAGCACACAAAAAAGCGTTAAAATTGGCTGAGCAAATTCAATGGTCTG  273
            ||||| |||| ||  |||  ||||||  || ||||| |||  ||||||   |||||||||
Db   620  ACTTCCGTGCGCAGCAGGCGCAGGCGCCGCCTACCAGCTGCTGCTGCTCGCTGGGACG  679

Qy   274  GGGCTTTTATCGTCGACACATTGGTTACAGGGCGTGTGGAACGAGACAAGCAAA  328
            |||  ||| |||| ||||    |||     ||| ||||||||||||||||||
Db   680  GTAGCTTTAGCCGTAGCCGGATATCGCGCTGCGGCGCGCGCCATCGAACGTGAAAAGGCTAA  734
```

RESULT 15
US-10-204-708-80
; Sequence 80, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8

; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 80
; LENGTH: 8961
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (5096)
; OTHER INFORMATION: n is a or g or c or t
US-10-204-708-80

  Query Match   4.1%; Score 67.6; DB 4; Length 8961;
  Best Local Similarity 45.2%; Pred. No. 1.6e-05;
  Matches 373; Conservative   2; Mismatches 437; Indels 14; Gaps 3;

```
Qy   332  GTTAGAAATCCTGTTGAATTAATTAGATAAAAAATATTGTACAGGGTAGAATTGTATTT  391
            |||   ||||| || | ||||  |||| || || |||
Db  5447  GTTTTATATTTTTATATATTAGTTTTATTTAAATGATTATTTATTATAATTATATTTTT  5506

Qy   392  TCCTAGGATTAGGATTTGTTGAGGCAACGTTTACGATTGCTCTGACAATAAATTAGAA  451
            ||   |||  ||||  ||||||||||        ||||    || ||    || ||||
Db  5507  TATTTATTTTAGTTTGATATTTTTTATAATAGTTAGAAATTTATTTTTTTTTTATATTATAA  5566

Qy   452  T-----TATTATTTTGTTACTTTATGAGGTTATATCAACTTATCGCACAATTTGTCAT  505
            |     |||  ||| || |||  ||||| |||||| |     |||| ||  ||||
Db  5567  TAACGGATATTCGGTTTTTTTTATATAATATTCGTAATAATTAGTTTTTTTATTTTTTAT  5626

Qy   506  CGTAGTATTATATTCTGCAATTCTCTTCATTAGAAGTATTATTATGAGAAAAATTATTAA  565
            || |||||||||| || |||  || || |     ||||||  ||| |
Db  5627  ATTGATATTTTATATAATAATTTTTTAGAGTTGTTTTTGTTTTTGTAGGTTGTTTTAGGGTT  5686

Qy   566  TATTTCTAGTATTGAGAGATACACTGAAAGTTTATTGTCTGTTTGTTGTTGTT  625
            |||  || ||| || |||  | |   |||| |||||     ||| ||
Db  5687  TAGAGGTTTATATTTTTAGTTTTTTTTTTAGTTTTAAGTAGTAAGTGATT  5746

Qy   626  GTTGTTTTCAAAATATAGAATTACAAGAATATAGTTGGCGCTTATTGCAATAAG  685
            |||  ||| ||||       ||  ||||  || | ||| ||  ||||||
Db  5747  TTTTTTTTATTATTAAAAAGATAATAATTTAATGATTTTTATTGTCGTTTTATAGAATG  5806

Qy   686  TATTGTGTTAATAATGTACATTATGCAGTATCAGTATACCAATCTGGATTGGACCTGTTAATTA  745
            |||||| ||||    |  ||     |||| ||||| |||||||| | |||  ||  |
Db  5807  GATATTTAATTTTTTTTTTTTTAATATTGGAAAGTTATTTTAATAATTTATTTTTT  5866

Qy   746  CTCACTTGCATTAAAGAAATTAAGAAATTAATGAGACTGGCTAACAATGACAGAAGCAA  805
            ||| || |||||| ||||  ||||||  |||||||  ||||| ||||||||||
Db  5867  ATTTATATTTGATATTTTAATAATTGTAAGGAGTTCGTGGTTAATTAAATATATTAT  5926

Qy   806  ATTCATATATCCATTGTTATTGGTTGGTTATTTTTTTATAATTTATTTTATTTTATTTT  864
            ||         |  |||   ||| | |  |||| | |||| | |||||   |||  |||
Db  5927  AATATTTTAATATATTTTTTTAGTGTTATATTTTTATAATGTTTTTATTTTTATTTTATTTT  5986

Qy   865  TCATAAAAAGAAAAGTATATAAACTTTCTTGGATTTTGACTTTATTTTTTTTTTATTTTTTTAGG  924
            ||        || ||  |    ||||      ||   ||||    ||||||  |||||||||||
Db  5987  TTATATTTTCGAGTTAGGTATTTTTAATATATTTTTATAATTTTTTTATTTTTTAGG  6046

Qy   925  TGATGTATGTTTTGTTCGAGCGTATACACAAAAATCCCATGAGCGTTTTATTCACCTA  984
            ||| |||| ||| |||      |    ||| ||||  |  |||||||      ||| |||
Db  6047  CGTTGTTAGTTTATTTAATATTTAATATTTTAGTAGTAGGGA------GGGTTAGGTAGGTTATTTA  6099

Qy   985  ACACTGTTATTCTCGATTAAAATCCAATTATTATTCGTTGGGTTATTATTTATAGGACGAA  1044
            |||||| ||| |  || |||||  |  |   | ||||  |||    |||  |||| |||
Db  6100  CGTTTTTTTTGTTTTCGTTTTTATTTTGTTATATTTTTTAATTTTTTTGTATATATTTTGGGTTAAAGTAGT  6159

Qy  1045  TTGTTCCTTATGAGATATTTCTTTATCTAATATTCCTCTTATCATAAATCTAAGCCTA  1104
```

Db   6160 ATTTTGGTAAAAAGTTTTTTGATGTTGGGATTTAGTTTTTTTATTATTATTATTTTTA 6219
                  |||| |||| ||| |||| ||| | ||| | |||| ||||| |||| ||| ||
Qy   1105 TGAAATCGGGCTCTCCGVAAAATTMARGAATATAATTTTAATTAAT 1150
                  ||||| |: | | || :| | || :| | || |||
Db   6220 GATATTCGGGGGTTGGGAGTTTATTTAGGGTTTATTTAGTTTTAT 6265

Search completed: December 23, 2004, 19:48:59
Job time : 156 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 23, 2004, 18:16:19 ; Search time 903 Seconds
                                            (without alignments)
                                            1011.434 Million cell updates/sec

Title: US-10-698-235-1
Perfect score: 1656
Sequence: 1 aagtaatatcacaataggg........ggtgttkggaataaataacg 1656

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 4105333 seqs, 2784095677 residues

Total number of hits satisfying chosen parameters: 8210666

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications NA:*
                1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
               10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
               11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
               12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
               13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
               14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
               15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
               16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
               17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
               18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
               19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
               20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
               21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	Length	ID	Description
1	292.4	17.7		1830121	14	US-10-329-960-1	Sequence 1, Appli
2	292.4	17.7		1830121	16	US-10-329-670-1	Sequence 1, Appli
3	292.4	17.7		1830121	18	US-10-158-865-1	Sequence 1, Appli
4	94	5.7		3673778	15	US-10-312-841-2	Sequence 2, Appli
5	89.4	5.4		8056	18	US-10-473-126-386	Sequence 386, App
6	87.2	5.3		19380	16	US-10-221-613-390	Sequence 390, App
7	86.6	5.2		8079	16	US-10-240-589C-122	Sequence 122, App
8	86.4	5.2		7814	16	US-10-221-714A-252	Sequence 252, App
9	85.8	5.2		3673778	15	US-10-312-841-1	Sequence 1, Appli
10	85.6	5.2		3683	18	US-10-473-126-339	Sequence 339, App
11	85	5.1		1105	18	US-10-425-115-173590	Sequence 173590,
12	84.6	5.1		1243	18	US-10-425-115-172717	Sequence 172717,

13	84.2	5.1		6255	15	US-10-311-455-933	Sequence 933, App
14	83.8	5.1	c	8056	18	US-10-473-126-240	Sequence 240, App
15	83.6	5.0		8776	16	US-10-257-166-150	Sequence 150, App
16	83.4	5.0		5822	15	US-10-311-455-1069	Sequence 1069, Ap
17	83.2	5.0		1062	18	US-10-425-115-120013	Sequence 120013,
18	83.2	5.0		7892	16	US-10-257-166-138	Sequence 138, App
19	82.8	5.0		6419	15	US-10-311-455-240	Sequence 240, App
20	82.8	5.0		113515	15	US-10-311-455-2147	Sequence 2147, Ap
21	82.2	5.0	c	8056	18	US-10-473-126-386	Sequence 386, App
22	81.8	4.9		6668	15	US-10-311-455-1670	Sequence 1670, Ap
23	81.4	4.9		17848	14	US-10-239-676-28	Sequence 28, Appl
24	81.4	4.9		17848	16	US-10-240-453-38	Sequence 38, Appl
25	81.4	4.9		17848	16	US-10-257-166-58	Sequence 58, Appl
26	81.2	4.9		3683	18	US-10-473-126-193	Sequence 193, App
27	81.2	4.9		14147	16	US-10-172-086-51	Sequence 51, Appl
28	81.2	4.9		14147	16	US-10-221-714A-469	Sequence 469, App
29	81.2	4.9		14147	17	US-10-311-507-39	Sequence 39, Appl
30	81.2	4.9		14147	18	US-10-480-846-51	Sequence 51, Appl
31	81	4.9		1081	18	US-10-425-115-16756	Sequence 16756, A
32	81	4.9	c	158001	16	US-10-211-179-11	GENERAL INFORMATI
33	80.6	4.9		921	18	US-10-425-115-38710	Sequence 38710, A
34	80.6	4.9		15548	15	US-10-311-455-2128	Sequence 2128, Ap
35	80.4	4.9		6292	16	US-10-221-714A-461	Sequence 461, App
36	80	4.8		7442	16	US-10-221-714A-409	Sequence 409, App
37	79.6	4.8		12405	14	US-10-239-676-35	Sequence 35, Appl
38	79.6	4.8		12405	15	US-10-240-453-43	Sequence 43, Appl
39	79.6	4.8		12405	15	US-10-221-613-101	Sequence 101, App
40	79.2	4.8		5979	14	US-10-239-676-17	Sequence 17, Appl
41	79.2	4.8		5979	15	US-10-240-453-25	Sequence 25, Appl
42	79.2	4.8		83391	17	US-10-433-793-123	Sequence 123, App
43	79	4.8		5452	15	US-10-311-455-1122	Sequence 1122, Ap
44	79	4.8		7498	15	US-10-311-455-230	Sequence 230, App
45	79	4.8		18683	15	US-10-311-455-286	Sequence 286, App

ALIGNMENTS

RESULT 1
US-10-329-960-1

; Sequence 1, Application US/10329960
; Publication No. US2003009277A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fragm
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,960
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40810)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature

LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)

Query Match   17.7%;   Score 292.4; DB 14;   Length 1830121;
Best Local Similarity   85.1%;   Pred. No. 5.6e-43;
Matches 336;   Conservative   1;   Mismatches 57;   Indels   1;   Gaps   1;

Qy   8 ATCACAATAGSGGATCCACGAGCTTCATTAGGTATCGTATTGGCTGCGAGGGATATCC   67
                  || |||||| |||||||||| ||||||||||||||| ||||| ||||||||||
Db   941919 ATATCACAATGGGATCTGCGAGCTTCT-TTAGTATCGTATTGGCTGCGAGGGATACCC 941977

Qy   68 AAAGGATTATCGCAAAGGCGATGAAATCAGCGGATTGCCTAAAAGTGCGGTCAAAAACGA   127
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   941978 AAAAGATTATCGCAAAGGCGATGAAATCAGCGGATTGCCTAAAAGTGCGGTCAAAAACGA 942037

Qy   128 GAAAGTTTTCTTAGCGGGTGTCGCAGAACAAGAAGGCAAGCTAGTCACAAACGGCGGTCG   187
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942038 GAAAGTTTTCTTAGCGGGTGTCGCAGAACAAGAAGGCAAGCTAGTCACAAACGGCGGTCG 942097

Qy   188 TGTACTTTGTGTGCGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA   247
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942098 TGTACTTTGTGTGCGACTGCGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA 942157

Qy   248 ATTGGCTGAGCAAATTCAATGCTCGGCGCGTTTTATCGTCGAGACATTGGTTACAGGGGC   307
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942158 ATTGGCTGAGCAAATTCAATGCTCGGGCGTTTTTATCGTCGAGACATTGGTTACAGGGGC 942217

Qy   308 TGTGGAACGAGAACAAGCAAAAATAGTTAGAAATCTTGTTGAAATTAATTAGATAAAAAAT   367
                |||||||||| ||| ||| |||| ||||||||||||||||| ||||||
Db   942218 TGTGGAACGAGAGTTGCAAAATAAATCCATATCATTATGATGAATTAATTTCATTTAAGT 942277

Qy   368 ATTGTACAGGGTAGAATTGTATTTTCCTAGGATTT   402
                ||||| | ||| ||| ||||||
Db   942278 GATGAAAATTTAGGAATGAAAATTTTAATTAAATTT 942312

RESULT 2
US-10-329-670-1
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)

Query Match   17.7%;   Score 292.4; DB 16; Length 1830121;
Best Local Similarity   85.1%;   Pred. No. 5.6e-43;
Matches 336; Conservative   1; Mismatches 57; Indels 1; Gaps 1;

```
Qy        8 ATCACAATAGSGGATCCACGAGCTTCTATTAGGTATCGTATTGGCTGCTGCAGAGGGATATCC 67
              ||||||| ||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   941919 ATATCACAATGGGATCTGCCGAGCTTCT--TTAGGTATCGTATTGGCTGCTGCAGAGGGATACCC 941977

Qy       68 AAACGGATTATCGCAAAGGCGATGAAATCAGCGGATTGCCTAAAAGTGCGGTCAAAAACGA 127
             |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   941978 AAAAGATTATCGCAAAGGCGATGAAATCAGCGGATTGCCTAAAAGTGCGGTCAAAAACGA 942037

Qy      128 GAAAGTTTCTTAGCGGGTGTCGCAGAACAAGAGGCAAGCTAGTCACAAACGGCGGTCG 187
             ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942038 GAAAGTTTTCTTAGCGGGTGTCGCAGAACAAGAGGCAAGCTAGTCACAAACGGCGGTCG 942097

Qy      188 TGTACTTTGTGTGACTGCGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA 247
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942098 TGTACTTTGTGTGACTGCGTTAGGCGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA 942157

Qy      248 ATTGGCTGCAGCAAATTCAATGGTCTCGGGCGTTTTATCGTCGAGACATTGGTTACAGGGC 307
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942158 ATTGGCTGCAGCAAATTCAATGGTCTCGGGCGTTTTATCGTCGAGACATTGGTTACAGGGC 942217

Qy      308 TGTGGAACCAGAACAAGCAAAATAGTTAGAAATCTTGTTGAATTTAATTAGATAAAAAAT 367
             ||||||||||||  ||||| ||||| || ||||    |||     |   |    |||||
Db   942218 TGTGGAACCAGAGAGTTGCAAAATAATCCATATCATTATGATGAATTTCATTTAAGT 942277

Qy      368 ATTGTACAGGGTAGAATTGTATTTCCTAGGATTT 402
              ||||   |||   ||||| ||||||||||||||
Db   942278 GATGAAAATTTAGGAATGAAATGAAATTTTAATTAAATTT 942312
```

RESULT 3

```
US-10-158-865-1
;  Sequence 1, Application US/10158865
;  Publication No. US20040203093A1
;  GENERAL INFORMATION:
;  APPLICANT: Fleischmann et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
;  TITLE OF INVENTION: Thereof, and Uses Thereof
;  FILE REFERENCE: PB186P2C1D1
;  CURRENT APPLICATION NUMBER: US/10/158,865
;  CURRENT FILING DATE: 2002-06-03
;  PRIOR APPLICATION NUMBER: US 09/557,884
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US 08/476,102
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/426,787
;  PRIOR FILING DATE: 1995-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: PatentIn version 3.1
;
;  SEQ ID NO 1
;    LENGTH: 1830121
;    TYPE: DNA
;    ORGANISM: Haemophilus influenzae
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (4747)..(4747)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (9921)..(9921)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (10150)..(10150)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (29298)..(29298)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (36543)..(36543)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (36551)..(36551)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (40808)..(40810)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (44416)..(44416)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (44905)..(44905)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (44975)..(44975)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (45593)..(45593)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (45732)..(45732)

;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (47036)..(47036)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (51334)..(51334)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (51602)..(51602)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (51786)..(51786)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (51805)..(51805)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (55369)..(55369)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (65309)..(65309)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (65313)..(65313)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (80024)..(80024)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (100091)..(100091)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (102696)..(102696)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (105121)..(105121)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (107248)..(107248)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (117136)..(117136)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (119750)..(119750)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (119924)..(119924)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (120038)..(120038)
;    OTHER INFORMATION: n equals a,t,c, or g
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (121344)..(121344)
;    OTHER INFORMATION: n equals a,t,c, or g
;
```

```
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (122167)..(122167)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (122336)..(122336)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (131340)..(131340)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (131360)..(131360)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (139910)..(139910)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (140398)..(140398)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (142750)..(142750)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (145058)..(145058)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (145171)..(145171)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (145942)..(145942)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (147197)..(147197)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (150841)..(150841)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (152500)..(152500)
        OTHER INFORMATION: n equals a,t,c, or g
        FEATURE:
        NAME/KEY: misc_feature
        LOCATION: (152530)..(152530)

Query Match         17.7%;      Score 292.4;   DB 18;   Length 1830121;
Best Local Similarity   85.1%;    Pred. No. 5.6e-43;
Matches  336;   Conservative   1;   Mismatches   57;   Indels   1;   Gaps   1;

Qy        8 ATCACAAATAGGGGATCCCACGAGCTTCTATTAGGTATCGTATTGGCTGCAGAGGGATATCC 67
            ||||||||| |  : ||||  | |||||||||||||||||||||||||||||||||||||||
Db   941919 ATATCACAATAGGGGATCTGCGGAGCTTCT-TTAGGTATCGTATTGGCTGCAGAGGGATACCC 941977

Qy       68 AAAGGATTATCGCAAAGGCCGATGAAATCAGCGGATTGCCTAAAAGTGCGGTCAAAAACGA 127
            ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db   941978 AAAGGATTATCGCAAAGGCCGATGAAATCAGCGGATTGCCTAAAAGTGCGGTCAAAAACGA 942037

Qy      128 GAAAGTTTCTTAGCGCGGTGTCGCAGAACAAGACAAGCAAGCTAGTCACAAACGGCGGTCG 187
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942038 GAAAGTTTCTTAGCGCGGTGTCGCAGAACAAGAACGCAAGCTAGTCACAAACGGCGGTCG 942097

Qy      188 TGTACTTTGTGTGACTGCGGTTAGGCGGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA 247
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942098 TGTACTTTGTGTGACTGCGGTTAGGCGGAAAGTGTATTTGAAGCACAACAAAAAGCGTTAAA 942157

Qy      248 ATTGGCTGAGCAAATTCAATGGCGTTTTATCGTCGCGTTTTATCGTCGCGAGACAATTGGT 307
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942158 ATTGGCTGAGCAAATTCAATGGCGTTTTATCGTCGCGAGACAATTGGTTACAGGGC 942217

Qy      308 TGTGAACGAGAACAAGCAAAATAGTTAGAAATCTGTGTGAATTTAATAGATAAAAAT 367
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   942218 TGTGGAACGAGAGTTGCAAAAATATCATATCATTATGATGAATTTCATTTAAGT 942277

Qy      368 ATTGTACAGGTAGAATTGTATTTTCCTAGGATTT 402
            ||||||||||||||   |||||| ||||||
Db   942278 GATGAAAAATTTAGGAATGAATGAAATTTTAATTAAAATTT 942312
```

RESULT 4
US-10-312-841-2
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
;   APPLICANT: Epigenomics AG
;   TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parameteen innerhalb des MHC
;   FILE REFERENCE: E01/1208/WO
;   CURRENT APPLICATION NUMBER: US/10/312,841
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
;   LENGTH: 3673778
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (379615)
US-10-312-841-2

Query Match   5.7%;   Score 94;   DB 15;   Length 3673778;
Best Local Similarity   46.3%;   Pred. No. 3.1e-06;
Matches 334;   Conservative   3;   Mismatches 384;   Indels   1;   Gaps   1;

```
Qy      422 GTTTACGATTGCTCTGCACAATAAATTAGAATTATTATTTTGTTACTTTATGAGGTTATA 481
            ||||                    ||||||||||||||||||||||||||||||||||||||
Db   257496 GTTTTTATTAGTTTTTTTATTATTTTTTTTTAATTTTTTTTTTTTTTTTTATTATTTTA 257555

Qy      482 TCAACTTATGCGACAATTGTCATCGTAGTATTATATTATCTGCAATTCTTCTTTCATTAGA 541
            |||     |||    |||||||     |||||    || |     |||||| || |
Db   257556 TTAATTTATTTATTTTTTAGTTAATTATGTAATTATATATATTTATTTTTTTTATTAT 257615

Qy      542 AGTTATTTATAGAAAATATTATTATTATTCTAGTATTGGAGAGATACACTGAAAGTTATTT 601
            ||||||||||||| |||| | ||||    |||||||    || |||   |||||||||||
Db   257616 TTATTATTTATGTTATTATTATACGTTTATTTAGTTAGTTATGTATGAAGGTTTATTATTTATTATT 257675

Qy      602 GTCTGTTTGTTTGTTTGTTTGTTTGTTTTTCAAAATATAGAATTACAAGAAATATT 661
            ||||||||||||||||||||||||||     |||||||||||||||| ||||||||
Db   257676 TGTTAATTTTTTTTTGTTTGTTGTTTAATTATTTTTTTTTAAATTATTTATTGTTTTTTATT 257735

Qy      662 AGTTGGCGCTTATTTGCAATAAGTATTGTGTTAATAATGTCAGTATACCA 721
            |||| || ||||||    ||||| ||||||||| |  |||||||||
Db   257736 TATTTATTTTTTTATTAGTTATTTTATTTATTTTTTTTTTTATT 257795

Qy      722 ATCTTGGATTGGACCTGTTAATTACTCACTGTATTATTTAAAGAAAATAATGAGATAACAA 781
            |||| |||| |||| |||  |||||||| ||  ||||||  ||||||||  |||| ||||||
Db   257796 ATTTATGTATATAAGGATGTATGTATTTATTGGTTATTATTTGGTTATTTAGGTTAT 257855

Qy      782 TGCTGGCTTAACAATGATGAGATAAAATCATATATCCATTGTATTTTTTGGTTTATTGAAGT 841
            |||                ||||    |||||||  |  |||||||||||||||||  |||||
Db   257856 TAATTAATTATTTTTTTTTTTTAATATATTTTTATAAATATTTATTTTTATTTATTTT 257915

Qy      842 TGCTCTGTTTTTAAGTTAAGTTTCATAAAAAGAAAGTATAAAACTTTCTTGGATTTT 901
            ||||          |||||||||||| ||||||||  |||||||  ||||||||||
Db   257916 TTATTCGATTAGTTATTATT-ATTTATAAGTTTATTTATTTTATTTTGTT 257974
```

```
Qy    655 GAATATTAGTTGGCGCTTATTGCAATAAGTATTGTTGTTAATAATGTACATTATGCAG 714
          |   |||||||||  |||| |  |||| ||  |||| |  |||||||    ||| ||||
Db   1172 TATTATTAATTATTTTGTAATTTTAATATTTTAATATATTTTTAATAATAAAT      1231

Qy    715 TATACCAATCTTGGATTGGACCTGTTAATTCTCGATTAAAATCCAATTATTTATC   774
          ||  |    |   ||| ||||  ||||   |||  |  ||| ||||||||| |
Db   1232 AATTTTTTGAATAAAAAAATGTTATTATATATATAAAAAATGTAATATT         1291

Qy    775 TAACAAATGCTGGCTTAACAATGATCAGATAAATTCATATATCCATTGTTATTGGTTAT 834
          ||||       ||||| |   |    ||| |  |||| ||  || |||| ||
Db   1292 TTTTAAAAATAATAAATAATATATAAAAAATTTATTATATAAAATAATTTAA      1351

Qy    835 TTGAAGTTGCTGTGTTTTAAGTTTATAAAAAATMARGAATATAATT           894
          |||   |     ||||  |||||||  |||| | | |:||
Db   1352 ATTTTAATTTTTATTTATTTTTTTTTTATAAAATAATAATTGATTTTTTTT       1411

Qy    895 GGATTTTTGACTTATTTTATGCTGTGATGATGTGTTTTGTTCGAGCGTATACAA     954
          |||      |   |||  ||||  ||||    |||  |||||||||
Db   1412 TATTTTTATTTTTTTTTTAAAAAAATAAAATATATTTTTTAATAAAATTAT       1471

Qy    955 CAAAATCCCATGAGCGTTTTATTCACCTAACACTGTTTATTCTCGATTAAAATCCAATT 1014
          |||        ||||  ||||||    ||| |  ||  ||||  ||||
Db   1472 ATTAAAATAATAATATTATTATTATA-TAAATATAAAATATGTTAAA           1530

Qy   1015 ATTTATCGTTGGGTTATTTATAGGACGAATTGTTCCTTATGGATATTTCTTTATCTA  1074
          ||   ||  ||||| ||  |||||   |||||||    |||||
Db   1531 AAAATTTTAATAAATTATTTTTTAAAAATATATTTAAATTTAATATTTTTAATTT     1590

Qy   1075 ATATTCCTCTTTATCATCAAATCTAAGCCTATGAAATCGGGCTCTCCGVAAATTMARGAA 1134
          ||||      |||||  |||    |||| || |||| |||||   |   |||:|:||
Db   1591 ATATTTTATTAATTTTAATATTTTTAAAAAATAAATAACTAAGAATT--GTAAAAATAAAAAA 1648

Qy   1135 TATAATTTTAATTAATGGGGAAAGTGCGCACCTCAAGTCATTTTAGTGTCTTTGGTTACG 1194
          |||||||     |||||||  |  |  ||| ||||||  |||
Db   1649 AATTATTTTTAATTAATATGATAAAATTTATTTTTATTAAAATATAATAATTTAA     1708

Qy   1195 GGAGAAAA--ACATCTCCTTTTTAGATAGCTTAAAATATAAATCAGGAGGTCTCTGTT- 1250
          ||| ||||  || ||  |||||   |||||||||    || ||
Db   1709 TATAATAATATATTTAATAATTTTAATATTTAATTAAAAAAATTAAAAAATTTTTTA      1768

Qy   1251 --GGTAAAACTTATTCCAGGACGAAAGCTAACAGCAATTTCTTTACCAATGTTTTTTAATG 1308
          |   ||      ||  |||||  ||  |||||||    ||||| |||||
Db   1769 AATTAAAAAAATTAAAATTAAAATATGTTTAATTTAATTAATATAATAAATAAAATT      1828

Qy   1309 CAATYCCTTAYCCAAATGGAATACAACAGATAGCTAAGGAGGATACGAATTTATTTAATT 1368
          ||  |  | | || |||  |   |||  ||| |  |||||| |  |||| |||||
Db   1829 TTGTTTTAAAAAAAAAAAAAAAAATAATTAATTTATTTTATTTTTTT           1888

Qy   1369 TAGGCGAAAGAGCAAGGCTTTCAGACATATTTTATTCAGCTCAAGCTAGGGATGATATGC 1428
          ||  ||||| |||||||||  ||   |  | ||||
Db   1889 TTTATTATAAAAATAAAAAATTTATAGGAGGATAAAATAATTATATAAATAAAATAAA     1948

Qy   1429 ATATGATCAAATTTTTAGGAGGGTAGAATTGTATTTCCTAGGATTTGTTGTTAGGG    1488
          ||||| |   ||||||||| ||  ||    |    |||   |  |||| |||||
Db   1949 AAAATATAATTTTTAAAAAATATTAATATTTAATTTTTATTTATTTAAAATTAATAAA     2008

Qy   1489 GGTATTCTTTAAGAGATTCAATGCCTGATAATAAATAATTACTTCCTGCTTTTAAAAATTA 1548
          ||||    |     | | ||| || |  ||  | |  | || |  |||||
Db   2009 TTTTAAATATATTTAAATTGTTTTTTTAAATGTTTTAAAAATGTTAT            2068

Qy   1549 ATTTAGATAAT                                                 1559
          ||   ||||||
Db   2069 TTAAAAATATT                                                 2079
```

RESULT 6
US-10-221-613-390

RESULT 5
US-10-473-126-386

```
Query Match       5.4%;    Score 89.4;  DB 18;  Length 8056;
Best Local Similarity  44.1%;    Pred. No. 2.1e-06;
Matches  587;  Conservative  3;  Mismatches 729;  Indels  12;  Gaps  5;
```

```
Qy    238 AAGGCGTTAAAATTGGCTGAGCAAATTCAATGTCTGGGCGTTTTATCGTCGAGACATTG 297
          ||   || |||||| | ||| |  |||||  |||||  ||||||||
Db    752 AATTTAAAAAATTTAATTAAAAATATAATATATAAATTAATTATATATTTAAAATAAATT 811

Qy    298 GTTACAGGGCTGTGAACGAGAACAGCAAAATAGTTAGAAATCTGTTGTGAATTTAATTA 357
          ||||| || | ||  | | ||  ||| ||||| | |     | |||| |||
Db    812 ATAAATGATTATTTAAAATATTTTATTTGTTTTTTTTATTTGTTTTTTTTATG      871

Qy    358 GATAAAAAATATTGTACAGGGTAGAATTGTATTTCCTAGGATTTAGGATTTGTTGTTAGGG 417
          |||||| |  | || ||||  |||||||  |  |||||  |||||||||| ||||| |||||
Db    872 TATAAAAATTTTGTAATTAAAATTTTTATTTATTTTTAAAATTAATAAATTAATAAAAA    931

Qy    418 CAACGTTACGATTGCTCTGACAATAAATTAGAATTATATTTTGTTACTTTATGAGGT    477
          ||||  |  || | |||  || ||| || | |   ||||| ||  |  |||| |||
Db    932 AATTTTTTTAAAAATAAAATTTTTATTTTTTTTAAATGAATAAAATAAAATATTATAA 991

Qy    478 TATATCAACTTATGCGACAATTTGTCATCGTAGTATTATATTCCTGCAATTCTTCCTTCAT 537
          ||||   || | || |   |||| || | |||| || | |||| | |||| |||||| |||
Db    992 TATGTAAAATAAAATAAATAATTTGTTTTTTTAATATTTAATTTTTTAAAATTTTTTAAAAT 1051

Qy    538 TAGAAGTTATTTATAGAAAATATTTAATATTTCTAGTATTGAGAGATACACTGAAAGTT 597
          ||  ||||| |  ||| |||| |||   |||| |||  |    | ||   ||| | ||
Db   1052 TATGTAAAATAAAAAATTTGTTTAATTTGTTTTTTTTTAATGTGTTTTTAAAATGTTAT 1111

Qy    598 ATTTGTCTGTTTGTTTGTTTGTTTGTTTTTTTTCAAATATAGA--ATTACAA        654
          |||| ||| |||| | || |||| |  |||||||  |||  |||  ||  ||||
Db   1112 ATATTTGTTTTTAATATTTTTTTTTGATATTTGAAAAATTTAAAAATTAATATATTTTTGAT 1171
```

```
Qy    902 TGACTTATTTTTATGCTGTGATGATGTATGTTTTGTTCGAGCGTATCAACAAATC     961
          |   ||||||||| | ||  |||  | ||  ||  |||| ||||
Db 257975 TTTTATTTATTTGTAATTTTGTTTATTTGTTAATTAGTTTTATTTTTATACGTATATT 258034

Qy    962 CCAATGAGCGTTTTATTCCACCTAACACTGTTTTTATTTCTCGATTAAAATCCAATTATTTATC 1021
          |    ||||  |||| | ||  | | || ||||   |    || ||
Db 258035 TATTTATTATTAATAATGTTTTAATATTGTGTATGTATGTATTTTTAGTTGTTATTT 258094

Qy   1022 GTTGGGTTATTTATTATAGGACGAATTGTTCCTTATGCTTATTTTCTTATCCTAATATTCC 1081
          ||| ||| |||||||   | ||   ||||   |||| |   ||| | | | | | ||||
Db 258095 ATTTATTTATTTATTATTTTTTATTTTTATATTTTTTTTTTTATTTTTTTTTA 258154

Qy   1082 TCTTTATCATAAATCTAAGCCTATGAAATCGGGCTCTCCGVAAAATTMARGAATATAATT 1141
          |||      ||| | |     ||||  || ||||   |     ||||| | |:||
Db 258155 AGTTAGTTAATATTTTTTGAATATTTTGTTGGATAAGGAGAATATAAAG 258214

Qy   1142 TT   1143
          ||
Db 258215 TT 258216
```

```
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
;   FILE REFERENCE: 5013.1004
;   CURRENT APPLICATION NUMBER: US/10/221,613
;   CURRENT FILING DATE: 2002-09-13
;   PRIOR APPLICATION NUMBER: PCT/EP01/02945
;       DE 10013847.00
;       DE 10019058.8
;       DE 10019173.8
;       DE 10032529.7
;       DE 10043826.1
;   PRIOR FILING DATE: 2001-03-15
;       2000-03-15
;       2000-04-06
;       2000-04-07
;       2000-06-30
;       2000-09-01
;   NUMBER OF SEQ ID NOS: 428
SEQ ID NO 390
;   LENGTH: 19380
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-390

Query Match             5.3%;  Score 87.2;  DB 16;  Length 19380;
Best Local Similarity  43.6%;  Pred. No. 7.6e-06;
Matches   386;  Conservative    1;  Mismatches  499;  Indels    0;  Gaps    0;
```

/ Mon Dec 27 09:35:02 2004   us-10-698-235-1.rnpb

```
Qy    632 TTTTTCAAAATATAGAATTACAAGAATATTAGTTGGGCTTTATTTGCAATAAGTA--TT 689
          |||| ||| ||  |  |||||||||||  | |||||||||||||||||||||||  ||
Db   5141 TTGTTCGGAATATTGTTTTTTTTTAGTAGTTTTTGGTTATTTATTTTTTTAAGTT 5200

Qy    690 GTTGTTAATAAATGACATTATGCAGTATACCAATCTTGGATTGGACCTGTTAATTACTTCA 749
          || ||| |||  ||||  |||| | ||||||||||||||||||||||||||||||||||||
Db   5201 TTTGTTTTATATTTATTTTGATTATTTTTGATTATTGGTAGTTTTTTTATTTTTTA 5260

Qy    750 CTTGCATTTAAAGAAATTAATGAGATAACAAAATGCTGGCTTAACAATGATAGATAAATC 809
          |||||| |||| |||| ||  || || ||||||||||||||||||||||||| ||||||
Db   5261 GTATTTTTTATTATTTTTTTTTTTTTTTTTTTTTTTTTTATATATATATAATTTA 5320

Qy    810 ATATATCCATTGTTATTGTTATTTGGTTTTGGTCGTGTGTTTTTTTTAAGTTTTTAAGTT 869
          |||||| || ||||||||||||||||||||||||||||||||||||| || |||||||
Db   5321 TTAATTTTGTGTATGTTTTGTATGTTTTTTTTTTTAGTAGAAGGTCGTTTGTTGTATTTTGT 5380

Qy    870 AAAAGAAAAGTATATAAAACTTCTGGATTTTTGACTTATTTTTTTTATGCTGTGATGATG 929
          |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5381 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 5440

Qy    930 TATGTTTTTGTTCGAGCGTATACAACAACAAAAATCCCATGAGCGTTTATTCACCTAACACT 989
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db   5441 TTTTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 5500

Qy    990 GTTATTCCTCGATTAAAATCCAATTATTTATGTTGGGTTATTTATAGGACGAATTGTT 1049
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5501 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 5560

Qy   1050 CCTTATGAGATATTTTCTTTATCCTAATATTCCTCTTTAT 1088
          |||| |||||||||||| |||||||||
Db   5561 TTTTTTTTTTTTTTTTTTTTTTTTTTT 5599
```

RESULT 8
US-10-221-714A-252
; Sequence 252, Application US/10221714A
; Publication No. US20040048254A1
; GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Diagnosis of Diseases Associated with
;   TITLE OF INVENTION: tumor suppressor genes and oncogenes
;   FILE REFERENCE: 5013.1005
;   CURRENT APPLICATION NUMBER: US/10/221,714A
;   CURRENT FILING DATE: 2003-01-21
;   PRIOR APPLICATION NUMBER: PCT/EP01/02955
;   PRIOR FILING DATE: 2001-03-15
;   PRIOR APPLICATION NUMBER: DE 10013847.0
;   PRIOR FILING DATE: 2000-03-15
;   PRIOR APPLICATION NUMBER: DE 10019058.8
;   PRIOR FILING DATE: 2000-04-06
;   PRIOR APPLICATION NUMBER: DE 10019173.8
;   PRIOR FILING DATE: 2000-04-07
;   PRIOR APPLICATION NUMBER: DE 10032529.7
;   PRIOR FILING DATE: 2000-06-30
;   PRIOR APPLICATION NUMBER: DE 10043826.1
;   PRIOR FILING DATE: 2000-09-01
;   NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 252
;   LENGTH: 7814
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-252

Query Match   5.2%; Score 86.4; DB 16; Length 7814;
Best Local Similarity 47.7%; Pred. No. 7.5e-06;
Matches 316; Conservative 0; Mismatches 341; Indels 5; Gaps 2;

```
Qy    429 ATTGCTCTGACAATAAATTAGAATTATTATTTTGTTACTTTATGAGGTTATATCAACTT 488
```

```
Db   3819 |||  |||  AATTTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTGTTT 3878

Qy    489 ATGCGACAATTGTCATCGTAGTATTATATATATCTCGCAATCCTCTTTCCAATTAGAAGTTATT 548
          ||||  ||  ||
Db   3879 ATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTAGTTTA 3938

Qy    549 TATAGAAAAATTAATATTCTAGTAGTAGATTGAGAGATACACTGA--AAGTTATTTGTCTG 606
          ||
Db   3939 TTTATTTATTTATTTATTTATTTCAAAATATAGAATTACAAGAATATTATTTTATTTTA 3998

Qy    607 TTTGTTGTTGTTGTTGTTGTTTTTCAAAATATAGAATTACAAGAATATTAGTG 666
          ||| || |||
Db   3999 TTTATTTATTTATTTAGTAGTTTTTCCGTTTTTAGGAGTTTTATGAGAATAGGAGTAGTATTTG 4058

Qy    667 GCGCTTATTGCAATAAGTAATGTTGTTGTAATAATGTACATTATGCAGTATACCAATCTT 726
          |||||| ||
Db   4059 ATTTTTTTTTGTTTTTGCGTTGTTTCGTATATGAATAAGTATGTTTTTTAATTTATTTATGTTGTT 4118

Qy    727 GGATTGGACCTGTAATTACTCACTTGCATTAAAGAAATTAATGACAATAACAAATGCTG 786
          |||||||||
Db   4119 GTAAAGATATAATTTATTTTTTTTTTTTATGGTTGAATAGATATTTATTTTGAAGTTGCTG 4178

Qy    787 GCTTAACAATGATAGATAAATCATATATAAATAATAGTGTTATTTATTTATTTCGTTGATTTG 846
          |||| ||
Db   4179 AATATAATAATAATATGTTATTTATTTATTTATTTATTTCGTTGATGGATATTTATGTTGATTTG 4238

Qy    847 TGTTTTAAGTTTAAGTTCATAAAAAGAAAAGTATATAAAATATATAAAATGTGTGTGTGT 906
          |||||
Db   4239 TATTTTTGTTATTATGAATAGTGTTGTGATAAAATATAAATGTAGTAGTATTTTTTGGTAT 4298

Qy    907 TTATTTTTATGCTGTGATGGATGATGTTTTGTTCGAGCGTATACAACAACAAAAATCCCATGTGG 966
          |||||||
Db   4299 TATTTATTTATTAGTTATT---TATTTATTTATTTATTTTTATTTATTTATTTATTTATTT 4355

Qy    967 AGCGTTTATTCACCTAACACTGTTTATTCTCGATTAAAATCCAATTATTTATTTATCGTGG 1026
          ||||||
Db   4356 ATTTATTATTTATTTATTTATTTTTTTTTATTTTGTTTTATTTATTTTTTAAGTTTATTTTT 4415

Qy   1027 GTTATTTTATAGGACGAATTGTTCCTTATGAGATATTTTCTTTATCCTAAATATTCCTCTTT 1086
          ||| ||
Db   4416 ATTTATTTATTTATTTATTTATTTATTTATTTATTTTATTTAGTTAGTTATTATTTTTATAAATTT 4475

Qy   1087 AT 1088
          ||
Db   4476 AT 4477
```

RESULT 9
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US2003018627A1
; GENERAL INFORMATION:
;   APPLICANT: Epigenomics AG
;   TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
;   FILE REFERENCE: E01/1208/WO
;   CURRENT APPLICATION NUMBER: US/10/312,841
;   CURRENT FILING DATE: 2002-12-30
;   NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
;   LENGTH: 3673778
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (3294164)
US-10-312-841-1

Query Match   5.2%; Score 85.8; DB 15; Length 3673778;
Best Local Similarity 48.8%; Pred. No. 0.0001;
Matches 298; Conservative 0; Mismatches 302; Indels 11; Gaps 2;

```
Qy  502  TCATCGTAGTATTATATTCTGCAATTCTTCTTCCATTAGAGAAGTTATTATAGAAAAATTAT  561
              |||  | ||||||| |||||||||||||||||||||||||
Db 2784  ATATATTTTTTTTATTTTTTTTTTAATTTTTGTTTTTTTAATTTTTTATTTTGTTTT 2843

Qy  562  TTAAATATTCTAGTATTGAGAGATACACTGAAAGTATTGTCTGTTGTTGTTTGTT  621
              |||||||  ||||  ||||||||| ||||||||||||
Db 2844  TTTTTTTTTATTTGTTTTTGGTTTGTTTTTTTGTTTTTTTTTTATTTTTT 2903

Qy  622  GTTTGTTGTTTTTCAAAATATAGAATTACAAGAGATAATTAGTTGGGCTTATTCGAA  681
              ||||||||||||||
Db 2904  TTTTTTTTGTTTTTTTGTTTTTATTTTATTTTATTTTTGTTGTTTTTGTTA 2963

Qy  682  TAAGTATTGTGTTAATAATGTACATTATGCAGTATACCAATCCTGGATTGGACCTGTTA  741
              ||||||||||| ||||||||||
Db 2964  TTTTTATT-TTTTTTTTTGTTTTTGTTTTTGTTTTTTTATTTTTTTTTTTTTTG 3022

Qy  742  ATTACTCACTGCATTAAAGAAATTAATGAGATAACAAATGCTGGCTTAACAATGATAG  801
              ||||  |||||||||||||
Db 3023  TTTTTTGTTTTTATTTTTATGTTTTTTATTTTTTTGTTTTTTTTATTTGATTT 3082

Qy  802  ATAAAATTCATATATCCATTGTTATTTGGTTATTGAAGTTGCTGTGTTTTTTTTTTAAGTTTAA  861
              ||||||||||||||
Db 3083  TTTATTTATTTTGTGTTATTTTATTTGTTTTTTTGTTTTTTGGTTTTTT 3142

Qy  862  GTTTCATAAAAGAAAGTATATAAACTTCTTCTTGGATTTTTGACTTATTTTTTATGCTG  921
              ||||||||||||||
Db 3143  ATTTTTGTTTTTTTTTTTTGTATGTTTTTTTTATTTTTTGTTTTTTTATGTT 3202

Qy  922  TGATGATGTATGTTTTTGTTCGAGCGTATACAACAAAATCCCAT-----GAGCGTTTTAT  976
              ||||
Db 3203  TATTTTTTTTGTTTTTGTTTTTTATGTTTATTTTTTTTATTTTTGTTTTTT 3262

Qy  977  TTCACCTAACACTGTTATTCTCGATTAAAATCCAATTATTATTATCGTTGGGGTTATTTAT 1036
              ||||||||||||||
Db 3263  TTTTTTTTATGTTTATTTTTTTTTATTTTTGTTTTTTTTATTTAATTTTTT 3322

Qy 1037  AGGACGAATTGTTCCTTATGAGATATTTTCTTCTATTCCTCTTTAT 1088
              ||||||||||||||
Db 3323  TGTTTTTTTTTTTTGTTTATTTTTTTATTTTTTTATTTTGTTTT 3374
```

```
Qy  335  AGAAATCTTGTTGACAATTAATTAGATAAAAAAAATATTGTACAGGGTAGAATTGTATTTTCC  394
              |||||||  |||||||||||||
Db 1058460  AGTTATTTAGTTAAAGTTATTGGGTGATAGATTTGGGAATAGAATTGGATTTTTAGATT 1058519

Qy  395  TAGGATTAGGATTTGTTAGGGCAACGTTACGATTGCTCTGACAATAAATTAGAATTA  454
              ||||||||||| ||||||||||
Db 1058520  TTTAGTTAGGGTTTTTTTTTTATTGTATTTGGGGGTTTTTTGATTTTTTTTTTTTTTT 1058579

Qy  455  TTAATTTTGTTACTTATGAGGTTATATCGACAATTGTCATCGTAGTATT  514
              |||||||||  |||
Db 1058580  TTTATTTTGTTAATTTAGTTTTTTGATTTTTTTATGGAAGGTTTTTTAATTTTTT 1058639

Qy  515  ATATTCTGCAATTCTTCTTCCATTAGAAGTTGATTATATAGAAAAATTAATAATTCTAG  574
              |||||||||||||| ||||||
Db 1058640  ATTTTTTTTTTTATATATTAGTTTTTAATATTATTTT-TAGTTTGAG 1058698

Qy  575  TATTGAGAGATACACTGAAAGTTATTTGTCTGTTGTTTGTT---------GTT  624
              |||||||||||||
Db 1058699  TATTTATTTTTATTTTGTTTTTGTTTTGTTGTTTAGATGTTT 1058758

Qy  625  TGTTTGTTTTTGTCGTTAAAATATGACAAGAATATTACAAGAATATAGTTGGGCGCTTTATTTGCAATAA  684
              |||||||||||||
Db 1058759  TTTTTATTTTGTCGTTAAAATGAGTTTTTATAGATGTGTACGTGTTGTTTTTGGTTTTAGA 1058818

Qy  685  GTATTGTTGTTAATAATGTACATTATGCAGTATACCAATCTTGGATTGGACCTGTTAAT  744
              ||||||||||||||
Db 1058819  ATTTTAGTGGTTTTTTTATTTTTTAGGTTTTTTTAGAAATATATTTAAAATTTTATATGGTGTATT 1058878

Qy  745  ACTCACTGCATTAAAGAAATAATGAGATAACAACAAATGCTGGCTTAACAATGATAGATA  804
              ||||||||||||||
Db 1058879  AGTTAGTTAGGTGTTATATAAATATTTTAGACGGGGTGGTTTAAATAATAGAAGT 1058938

Qy  805  AATTCATATCCATTGTTATTTGGTTTATTTGAAGTGCTGTGTTTTTTAAGTT  864
              ||| |||
Db 1058939  TATTTTTATAGTTTGGAAGTTTGAGAAGTTTAAGATTAAAAGTGTTGGTAAGGTTAGTTT 1058998

Qy  865  TCATAAAAAGAAAGTATATAAACTTTCTTGGACTTTTGACTTATTTTTTATGCTGTGA  924
              |||||||
Db 1058999  TATTTAGATTTTTAGAGTTTTTGTTTTTATTTTTTTATTTTTTATTTTTTT 1059058

Qy  925  TGATGTATGT  935
              |||
Db 1059059  TTATTTTTTTT 1059069
```

```
Qy  442  TAAATTAGAATTATATTTTGTTACTTATGAGGTTATATCAACTTATGCGACAATTTG  501
              ||||||||||||||
Db 2724  TATTTTTTTTTTTTTTTTGTTTGTTTTTTTTTTTTTTTTTAATTTTTTAATTTTTGTTTTTTTT 2783
```

```
Db   393 ||||| |||  |||||  ||||||||
         TTTTTGTGTTTTTGGTGTATTATATTGGTATTATATTGTTCTTCTTTCTCAATTTGTAAT 452

Qy   402 TAGGATTTTGTTAGGGCAACGTTACGATTGCTCGACAATAAATTAGAATTATTATTTT 461
         |||||         |||||||||||||||||||||||||||||||||||
Db   453 GTTCTTTTTTTTAAGGTTTTTGTTCTTTATTTGTTCTTTTATTTGTTATNTATTATTATTTATTAT 512

Qy   462 TGTTACTTTATGAGGTTATATCAACTATGCGACAAATTGTCATCGTAGTATTATATTCT 521
         ||  ||  |||   |||||||||||||||||||||||||||||||||
Db   513 --TTAATNTAATATTTTTTTTATTTATAAGTGTATTCTTTTTTTTAATTTTTTCT 570

Qy   522 GCAATTCTTCTTCATTAGAAGTTATTATTAATATTAATAATATTTCTAGTATTGAG 581
         ||||||||    |||| ||||
Db   571 TGTTTTCTTTTTCTTTTTTTAATTCCTTTTTT-----TTAATNTTCCTTTTTTATTTAT 624

Qy   582 AGATACACTGAAAGTTATTTGTCTCGTTGTTGTTGTTGTTTTTTCAAAA 641
         |||||||||||||||||||||||||||||||||||
Db   625 ATATAGAAATTCTTTTNTTTTTTTTTTTTNTATTTTTTTNTATTTAA 684

Qy   642 TATAGAATTACAAGAATATTAGTTGGCGCTTATTGCAATAAGTATTGTGTTAATAAT 701
         ||||||||||||||||||||||||||||||||||||||
Db   685 TTNATTTTTTTGTTTTTTTTNTTAATTTTTATTTTTTTCTTGTTCTTTTATTTTT 744

Qy   702 GTACATTATGCGAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTTGCATTTAAA 761
         |||||||||||||||||||||||||||||||||||||
Db   745 TTAATTTTTTCTATATTTTTTTACTTTTTTATTTTTATTTTTATTATTTA 804

Qy   762 GAAATTAATGAGATAACAAAATGCTCGCTTAACAATGATAGATAAATTCATATATCCATTG 821
         ||||||||||||||||||||||||||||||||||
Db   805 TTTTTTTTTTCTATATTTTTTATTTTTTTTTTTAATTTTTTAATTTTTTTTTTTT 864

Qy   822 TTATTTGGTTTATTTGAAGTTGCTGTGTTTTTAAGTTCATAAAAAAGAAAAGTA 881
         |||||||||||||||||||||||||||||
Db   865 TTTTTATTTTTTTTATTTTTTTTTTTATTTTTTTTATTTTTTTTTTTTTTTTT 924

Qy   882 TATAAACTTTCTTGGATTTTGACTTTATTTTTATGCTGTGATGATGTATGTTTTTGTT 941
         |||||||| |||||||| |||||||||
Db   925 TTTTTTTTTTTTATTTTTTTTTTTTTTTTTGTTTTTTATTTTGTTTTTTTTTTTT 984

Qy   942 CGAGCGTATACAACAAAATCCCATGAGCGTTTATTCACCTAACACTGTTTATTCTCGA 1001
         |||||||||||||||||||||||||||||||
Db   985 TTTTATTTTTTTATTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATT 1044

Qy  1002 TTAAAATCCAATTATTATTATCGTTGGGTTATTTTATAGGACGAATTGTTCCTATGAGATA 1061
         ||||| |||||| || | ||||||||
Db  1045 TTTTATTTATTTTTTTTTTTTTTTTTTTTTTTTNTTTGTNTNTTTTTTTTTTT 1104

Qy  1062 TTTTCTTTATCTAAATATTCCTCTTTAT 1088
         ||||| |||||  |||| |||
Db  1105 TTTTTTTTTTNTTTTTTTTTTTTTTTNNT 1131
```

RESULT 13
US-10-311-455-933
; Sequence 933, Application US/10311455
; Publication No. US2003014360A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 933

```
Qy   342 TTGTTGAATTTAATTAGATAAAAAAATATTGTACAGGGTAGAATTGTATTTCCTAGGATT 401
         ||||||    ||||||
Db   509 TTTTTGTGATTAAATTTGTTTTTTAGAAAGAATATTTGTAGTTTTTTTTTGTTTTTTTT 568

Qy   402 TAGGATTTTGTTAGGGCAACGTTACGATTGCTCGACAATAAATTAGAATTATTATTTT 461
         |||
Db   569 TTTTTTTGTTATTTTTTCATTTTTTGTTTTTTTTTTTATTTTATATCTTTTTTTTT 628

Qy   462 TGTTACTTTATGAGGTTATATCAACTATGCGACAAATTGTCATCGTAGTATTATATTCT 521
         ||| ||||
Db   629 TTTTTTTTGTTCTTTTATTCTTTTTTTTTTTGATATTTTTATTTTTTTTTTCTCTTTT 688

Qy   522 GCAATTCTTCTTCATTAGAAGTTATTATTAATATTAATAATATTTCTAGTATTGAG 581
         |||||||||||||||
Db   689 TTGAATTTCTTATTTCTTTTATTGTTTTTATTCTTTTTTTGTTTTTTTTCTTTTTTTATATAT 748

Qy   582 AGATACACTGAAAGTTATTTGTCTCGTTGTTGTTGTTGTTTTTTCAAAA 641
         |||||||||||
Db   749 TTTTTTTTTTGATTCTGTTTTTTTATTTTTTTATTTTTTTTTTTTATAATG 808

Qy   642 TATAGAATTACAAGAATATTAGTTGGCGCTTATTGCAATAAGTATTGTGTTAATAAT 701
         ||| ||
Db   809 TTTTTTTCATATTTATATGCTTATTTTTAGTTTTTTTTTAGTTTTTTTTTATATA 868

Qy   702 GTACATTATGCGAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTTGCATTTAAA 761
         |||||| |||||
Db   869 TTTTTTTTTTTTTTTTTTAATTTTTTCTTTTTTTTATTTTCATTTTTTTT 928

Qy   762 GAAATTAATGAGATAACAAAATGCTCGCTTAACAATGATAGATAAATTCATATATCCATTG 821
         ||||||||||||||
Db   929 TTTTTTTTTTTAAATTTTTCTTTTTTAATATTTTTTAATCATATTTTTTTATATATTT 988

Qy   822 TTATTTGGTTTATTTGAAGTTGCTGTGTTTTTAAGTTTCATAAAAAGAAAAGTA 881
         ||||||| ||||
Db   989 TTTTTTTTTAATTGTTAAATTGTTTTTTTTTATCTTTTTTTTATCTTTTTAAATTTTT 1048

Qy   882 TATAAACTTTCTTGGATTTTGACTTTATTTTTATGCTGTGATGATGTATGTTTTT 938
         |||||||| ||||| | |||||
Db  1049 TAGGAATTTTTTTATTTTTTTGTTTTTTTTTTTTTTGTTTTTTTTTTTTTTT 1105
```

RESULT 12
US-10-425-115-172717
; Sequence 172717, Application US/10425115
; Publication No. US2004021472A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 172717
; LENGTH: 1243
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1243)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_89100C.1
US-10-425-115-172717

Query Match   5.1%; Score 84.6; DB 18; Length 1243;
Best Local Similarity 45.9%; Pred. No. 8e-06;
Matches 343; Conservative   0; Mismatches 396; Indels   8; Gaps   2;

Qy   342 TTGTTGAATTTAATTAGATAAAAAAATATTGTACAGGGTAGAATTGTATTTCCTAGGATT 401

```
;   LENGTH: 6255
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-933

Query Match        5.1%;   Score 84.2;  DB 15;  Length 6255;
Best Local Similarity  45.1%;  Pred. No. 1.8e-05;
Matches 349;  Conservative 0;  Mismatches 423;  Indels 1;  Gaps 1;

Qy    326 AAAATAGTTAGAAATCTTGTTGAATTAATTAGATAAAAAATATTGTACAGGGTAGAATT 385
              ||| || |   ||||| || | || | |||| |  | |||
Db   5399 AGAAAAAGGTGTATTTTAAATGGAATTTAAGTATTAGAGGGTGTTAAAAATAGGAATTTTA 5458

Qy    386 GTATTTCCTAGGATTAGGATTTGTTAGGGCAACGTTTACGATTGCTCTGACAATAAA 445
              ||  ||  | |||||| || | |  |||| || | |||
Db   5459 AGTTATTTTTGTATAATTTATTTTTAAAGTTAGTTAGTATATATTTTAAATTTTTTTTAAA 5518

Qy    446 TTAGAATAATTAATTTTGTTACTTTATGAGGTATATCAACTTATGCGCACAATTTGTCAT 505
              || || | |  |||| ||  |||  | || || || ||| | | ||
Db   5519 ATAAAGTTGTTTATTTTAAAATAAAATAAATTATTTATATTTTATAGAAAAATTAAAATTTT 5578

Qy    506 CGTAGTATTATATTCTGCAATTCTTCTTCATTAGAAGTTATTTATAGAAAATTATTTAA 565
              || || || | | ||  |  | | | || |  |||| | || |  |
Db   5579 TTTAATTTATTTTATTTGTATTAGTATTAGAGGAAAAGTATTTGATTTTTTTTGGA 5638

Qy    566 TATTTCTAGTATTGAGAGATACACTGAAAGTTATTGTCGTTGTTGTTGTTGTTT 625
              |||    | |  | ||| |  | | || | ||||||||||||
Db   5639 GTTTTTTTTTTTTCGTTTTTATCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 5698

Qy    626 GTTTGTTTTTTCAAAATATAGAATTACAAGAAATAGTTGGCGCTTATTTGCAATAAG 685
              | |  ||| ||||||||| || |  |  ||| |||||
Db   5699 ATGAATTTAATAGGTTATTAAAAATAAAAAATTATTAAAATAATTTATAAAATATTATT 5758

Qy    686 TATTGTTGTTAATAATGTACATTATGCAGTATAACCAATCTGGATTGGACCTGTTAATTA 745
              ||| |  | |  ||| ||| || | || || ||||||||||
Db   5759 TTTAATGAATGTTTCGAAATTGAATTGGTTATTAGTTTTATTACGTGATTTTAT -AGTGAT 5818

Qy    746 CTCACTTGCATTAAAGAAAATTAATGAGATAAACAAATGCTGGCTTAACAATGATAGATGA 805
              || || || |  |  | |  | | || |  || | || ||| ||
Db   5819 TATAAAAGGATAGAAGTAGAAATAATAAATAAGAATAAGTAAGTATAATGTTTGT 5878

Qy    806 ATTCATATATCCATTGTTATTTGGTTTTATTTGAAGTTGCTGTGTTTTAAGTTT 865
              | || |  || || || | | || |||  | ||||| | | |||
Db   5879 TTTATTAGTAAATAAAAATTTGTATGTTCGGTAGTTTTGTAGTGTATGGAATATATTT 5938

Qy    866 CATAAAAAGAAAAGTATATAAACTTCCTGGATTTTCTTGGAATTTTTATTTTTTGACTGTGAT 925
              | || | || | |||  | | || | || | ||||
Db   5939 ATATAAAAAAAATAATTTTAGAGTTTTATATAAAAAGATATTTTTATTTTTTATTTTAT 5997

Qy    926 GATGTATGTTTTGTTCGAGGCGTATACACAAATCCCATGAGCGTTTTATTTATTTTATA 985
              |  | || || || |  | | ||||||
Db   5998 AGATTTTTTTTTTGTTTATATAATTTAGAGTTTATATTTATTTTTATTTAATTAATA 6057

Qy    986 CACTGTTTATTCTCGATTAAAATCCAATTATTATTATTCGTTGGGTTATTTTATTAGGACGAAT 1045
              || || | || | | || | ||| | || |  |||| || | |
Db   6058 TAGTATTAATTATATAGTTATTTTATTTTTATTTTATTTAATTTTTTAAAATAA 6117

Qy   1046 TGTTCCTTATGAGATATTTCTTTATTCATCTTCTCTAATAATCCTCTTAATCATAAATCTA 1098
              ||  | ||||| | |||  ||||
Db   6118 AATTAATTTAGTTGTTATTAATTAAAGAATGTTTTATTAATTTTGTTTTAATTATA 6170
```

```
;   CURRENT APPLICATION NUMBER: US/10/473,126
;   CURRENT FILING DATE: 2003-09-26
;   NUMBER OF SEQ ID NOS: 1258
;   SEQ ID NO 240
;   LENGTH: 8056
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-473-126-240

Query Match        5.1%;   Score 83.8;  DB 18;  Length 8056;
Best Local Similarity  46.0%;  Pred. No. 2.3e-05;
Matches 562;  Conservative 4;  Mismatches 638;  Indels 19;  Gaps 8;

Qy    443 AAATTAGAATTATATTTTGTTACTTATGAGGTATATCAACTTATGCCGACAATTTGT 502
              ||||||  |||| || | |  || |  ||  | ||  |  ||| |
Db   2391 AAATTTAATAAATTATTAATAAAAATTTAAAATTTAAAAATTAAAATTAAT 2332

Qy    503 CATCGTAGTATTATATTCTCGCAATTCTCCTTCATTAGAGTTATTTATAGAAAAATATT 562
              |||| || || |  || | |  | | ||||  |  ||| |  ||| || |||| || |
Db   2331 TTAAATACGAAAATTAATAAACATATTTTAAAATAAAATTAAAAATTTAAAATTTT 2272

Qy    563 TAA---TATTTCTAGTATTGAGAGATACACTGAAAGT--TATTTGTCGTTGTTGTTGTTT 617
              |||   |||  | || || | ||| |  |  || |  |||  |  || |||||||||||
Db   2271 TTAATTTATTTTATTTTTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTT 2212

Qy    618 GTTTGTTTGTTTTTCAAAATATAGAATTACAAGAAATAGAATTACAAGAATATATATTTATTT 677
              |  | |  | || | ||||| |||  |  |  ||  ||| | ||||||||||
Db   2211 TATTTTTTTTTCGTTTTTTTTAAATTTTTTAAATTTTATTAAAAATTATATAAATTTATTT 2152

Qy    678 GCAATAAGTATTGTTGTTAATAATGTACATTATGCAGTATACCAATCTTGATTGGACCT 737
              ||||||| |||| | | |  |||| ||| || | || || ||||||||||||||
Db   2151 TTTATTTTTAATAAAAATAAAAAATTATTAAAATAATTTATAATATTATATTATT 2092

Qy    738 GTTAATTACTCACTTGCATTAAAGAAAATTAATGAGATAACAAATGCTGGCTTAACAATG 797
              ||| || || |  |  | || ||| |  |  | ||   |  || || |||  | || ||
Db   2091 TATATATTTAAAATATTTTTTTAATAAAATTTTTTTAAAAATAAA--AATATTTTAAAAAT 2035

Qy    798 ATAGATAAAATTCATATATCCATTGTTATTTGGTTTTATTTGAAGTTGCTGTGTTTTTAAGT 857
              || ||| || || | |||  | ||  | || ||| | || ||||  | |  |||| ||
Db   2034 ATATTTTTATTTAAATATTTTTAAATTTTAAATATTTTAAAATTAATAATATTAAT 1975

Qy    858 TTAAGTTTCATAAAAAGAAAAGTATATAAAACTTTCTTGGATTTTGACTTTTTATTTTTTAT 917
              |||||| | || | || | |||  | | || ||  | | || |||| ||||||||||
Db   1974 ATTTTATTTTAAATAAAATTATATTTTTTTA-TTTATATTATAAATTTAATTATTTTTTTTT 1916

Qy    918 GCTGTGATGATGTATGTTTTGTTCGAGGCGTATACACACAAATCCCATGAGCGTTTTATT 977
              | | || || || || ||  | | | || ||||
Db   1915 TTTATAAAATGTTTTTTTTTTTTTTTTAAAACGAAAATTTATTTTTATTTTTTAATTTTTT 1856

Qy    978 TCACCTAACACTGTTTATTCTCGATTAAAATCCAATTATTTCTTTATTCTCTTTATCATAAA 1037
              | |  || || || | || | | || | ||| |||  |  |||| ||
Db   1855 TTTTTTTTTTTTTTTTTTTTTTTTTAAAACGAAATTTATTTTTAAAATATAATA 1796

Qy   1038 ---GGACGAATTGTTCCTTATGAGATATTTCTTTATTCATCTTTTATCTAATAA 1094
                 | |||  | ||  | ||||| | |||  ||||  || | | ||
Db   1795 ACGTAATTAATTTAATTTTTAATTTTTTAATTTTTAAAAAATTTTTTTAAATAA 1736

Qy   1095 TCTAAGCCTATGAAATCGGGCTCTCCGVAAAATTWARGAATATAATTTAATGGGG 1154
              | |||| || | || |  || ||  || |  || | | ||
Db   1735 TTAAAATAAATAAATATATTAAATTTTTTATAATTTTTATAAAAATAA 1676

Qy   1155 GAAAGTGCGACCTCAAGTCATTTAGTGCTTTGGTTACGGGGAGAAAAACATCTCCTTTT 1214
              ||||| |  |  ||  ||| || || | |  ||||| || |||||||
Db   1675 AATTTTATCGTATTATTAAATAATTTTTTTAATTTTTACGGTTCGTTTTATTTTCGATTT 1616

Qy   1215 TTAGTAGCTTAAAATAAATAAATCAGGAGCTCTTGTTGGTAAAACTTATTCAGGAGGAAAG 1274
              ||||   ||| |  | |||||||  |  || | || ||  || |
Db   1615 TTTAAAATAATTATAAAAATTAAATAAATAAAAAT--ATTAAATTTAAAATAATTTTTT 1558

Qy   1275 CTAACAGCAATTCTTTACCAATGTTTTTAATGCAATYCCTTAYCCAAATGGAATACAA 1334
              || |||||||||||    ||  |||| || ||  ||  | | ||||||
Db   1275 ... 1
```

RESULT 14
US-10-473-126-240/c
; Sequence 240, Application US/10473126
; Publication No. US20040234973A1
GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:

```
Db   1557 AAAAAAATATATTTTATTTAAAAATTTTTTAACGTATTT--ATATTTTATATATTATATAA 1500

Qy   1335 CAGATAGCTAAAGGAGATACGAATTTATTTAATTTAGGGAAAAGAGCAAGGCTTTCAGACA 1394
          |||||        | |||||       | |||||                      |||
Db   1499 TATTAATTTAATATAATTAATTTTTTAATATTTTTATTAAAAAAATAATATTTTATTTTT 1440

Qy   1395 TATTTTTATTCAGCTCAAGCTAGGGATGATATGCATAT--GATCAATTTTTAGGAGGA 1451
          |||||||       ||||| |  |                | |||||||
Db   1439 TTTTTTAAAAAAAAAAATAAAAAAAATAAAAATCGATATTATTTTTTATAAAAAAA 1380

Qy   1452 GCTTGGATTGATGATATCGTTTCCGAGATAATCGTATTCTTTAAGAGATTCAATG 1511
          |||||        | ||||         ||||| |||     | |||
Db   1379 AAATAAATTTTTTAATAAAATTTAAAAAATTAAATTTTTTATTAAAAAATTATAATAAA 1320

Qy   1512 CCTGATAATAAATTACTTCCTGCTTTAAAAATATTAATTTAGATAAGATGTTATCATTTT 1571
          |||||      |||||||        ||||| |  |||
Db   1319 TTTTTTATTAATTTATTAATTTTAAAAAAATAATATTACGTTTTTATATATTAATAATTAC 1260

Qy   1572 GTTGTTTACATCATAGAGGGAGTCATATTCCCTATGGGCATTATTAGGAATGAAAAAGN 1631
          || |||||||||       |  ||  |       || |||||
Db   1259 GTAACGTTTTTTTTTATTCGAAAAATATTTTTATTTATTTAAAAAAAATTAAATTAAATAT 1200

Qy   1632 AGNAAGGTGTTKGGAARAAATAA 1654
          ||  ||   || | |||||||||
Db   1199 ATTAAAATTTTAAAATAARATTAA 1177

RESULT 15
US-10-257-166-150
; Sequence 150, Application US/10257166
; Publication No. US2004002330A1
; GENERAL INFORMATION:
;   APPLICANT: OLEK, Alexander
;   APPLICANT: PIEPENBROCK, Christian
;   APPLICANT: BERLIN, Kurt
;   TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
;   TITLE OF INVENTION: Genes Implicated in Pharmacogenomics
;   FILE REFERENCE: 5013.1011
;   CURRENT APPLICATION NUMBER: US/10/257,166
;   CURRENT FILING DATE: 2002-10-07
;   PRIOR APPLICATION NUMBER: PCT/EP01/07470
;     DE 10032529.7                    2000-06-30
;     DE 10043826.1                    2000-09-01
;   PRIOR FILING DATE: 2001-06-29
;   NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 150
;   LENGTH: 8776
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-150

  Query Match    5.0%;    Score 83.6;  DB 16;  Length 8776;
  Best Local Similarity  47.0%;  Pred. No. 2.6e-05;
  Matches  359;  Conservative  0;  Mismatches 399;  Indels  6;  Gaps  3;

Qy    339 ATCTTGTTGAATTAATTAGATAAAAAATATTGTACAGGGTAGAATTGTATTTTCCTAGG 398
          ||||||||      ||       ||  |     | || |||       ||||||||
Db   6685 ATATTTTTATTTTTTGTTTTAGATGTTTATTGTTATTGGTATGTATGTAGAAATGTTATTTTTTTG 6744

Qy    399 ATTTAGGATTTGTTAGGGCAAGTTTACCGATTGCTCTCGACAATAAATTAGAATTATTAT 458
          |||||||||||||     ||  |||       || |   | | |||
Db   6745 TATGTTGATTTGATTTAGTAATATTATTAGAATTGTTTAATTAATATGAATTTAATAGTTTTTT 6804

Qy    459 TTTTGTTTACTTTATGGAGTTATATCAACTTATGCGACAATTGTCATCGTAGTATTATAT 518
          |||||     | |||| |       |  ||  ||||
Db   6805 AGTGGAGTTTTAAGTGTTTTTATATATTTTTATAATTATTTGGAAAATATTATTTTGGA 6864

Qy    519 TCTGCAATTCTTCTTTCATTAGAAGTTATTAATATAGAAAATTATTTAATATTTCTAGTATT 578
          ||        ||     | ||         |||||   |||||
```

Search completed: December 23, 2004, 21:59:10
Job time : 914 secs

```
Db   6865 AATATTTTTAAAATATTTTTAAAATATTATTTGGAAAATATTTGAAATATTTTTTGTT 6924

Qy    579 GAGAGATACACTGAAAGTTATTTGTCTGTTTGTTGTTGTTTTCA 638
          |        |||  |||
Db   6925 TTTTGGTATA---AAATGTATTTTAGGTTATTTTTTATTTGAATATTTTTAATTTTGGTTTTTGGA 6981

Qy    639 AAAATAGAATTACAAGAGATATTAGTTGGGCGCTTATTTGCAATAAGTAATGTATGTGTTAAT 698
          ||||||                  ||||| ||||       ||||| |||||
Db   6982 ATTAGTTATTTTTTAAAAAGTTTTGGTTTTTAGTGAAAAATGTATTAATAAGTTA 7041

Qy    699 AATGTACATTATGCAGTACCAATCCTTGGAYTTGGACCTGTTAATTACTCACTTGCATT- 757
          |||||        | |||       ||| ||||| |||
Db   7042 AGATTTGGTTATTAAGGTGTTATTGTTATTGAAATGTTATTGTTATTTAGTTTTATTA 7101

Qy    758 -TAAAGAAATTAATGAGATAACAAAATGCTGCTTAACAATGATAGATAAATTCATATATC 816
          |||||       ||||| |    |||||||
Db   7102 GTGGAGAGTAGGGAGTATATGTTTTTGTATATTAAATATAAATATAATATTATATATT 7161

Qy    817 CATTGTTATTTGGTTATTTGAAGTTGCTGTGTTTTAAGTTAAGTTTCATAAAAAGAA 876
          |||||       |  | ||       |||| ||||
Db   7162 TATGTTTATTTAATTATTTTTTGTTATTTTTGTTATTAATTATTTAATTATTTAT 7221

Qy    877 AAGTATATAAACTTCTTGGATTTTTGACTTTATTTTATGCGTGTGATGATGTATG-TT 935
          |||||       ||  |||||||     | ||||
Db   7222 TTATTTAGAGGGTATGAATATATATATGGATATATTTATTTTAATTTAATTTATACGGATT 7281

Qy    936 TTTGTTCGACGTATACAACAAAATCCCATGAGCGTTTTATTTCACCTAACACTGTTTAT 995
          ||||||       |  |             | |||  ||      | ||||
Db   7282 TTTGTTTAGTTTTTTTTTGTTTTGAATTTATGATTTTTTTTTGTAAAAGTGATATATTTGG 7341

Qy    996 TCTCGATTAAAATCCAATTATTATTCGTTGGGTTATTTTATAAATTTTTTTTATGTAGTGAAATTTTTATTT 1055
          ||         |   || |||||       |||       ||||||||
Db   7342 TTTTTAGTATATTTTATTTATTTTTTATTAATTTTTTATGTAGTAGAAATTTTTATTT 7401

Qy   1056 GAGATATTTTCTTTATCTAAATATTCCTCTTTATCATAAATCTAA 1099
          |||||          |        |||| |||  |||
Db   7402 ATTTTTTTTTTATTTATTATGGATGTTTTTTTTTTATTTTTTATGTTTTAAA 7445
```

This Page Blank (uspto)

```
               GenCore version 5.1.6
          Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:      December 23, 2004, 15:46:49 ; Search time 5397 Seconds
                                  (without alignments)
                            11181.063 Million cell updates/sec

Title:             US-10-698-235-1
Perfect score:     1656
Sequence:          1 aagtaatatcacaatagsgg.........ggtgttkggaaraaataacg 1656
Scoring table:     IDENTITY_NUC
                   Gapop 10.0 , Gapext 1.0
Searched:          32822875 seqs, 18219865908 residues
Total number of hits satisfying chosen parameters:    65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
                 1: gb_est1:*
                 2: gb_est2:*
                 3: gb_htc:*
                 4: gb_est3:*
                 5: gb_est4:*
                 6: gb_est5:*
                 7: gb_est6:*
                 8: gb_gss1:*
                 9: gb_gss2:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                            SUMMARIES
```

	Result No.	Score	% Query Match	Length	DB	ID	Description		
	1	107.8	6.5	1254	9	AG349719	AG349719	Mus	muscu
	2	103.8	6.3	2157	9	CL081966	CL081966	CH216-169	
	3	103.2	6.2	1277	8	CC253231	CC253231	CH261-180	
c	4	102.8	6.2	1536	9	CL078538	CL078538	CH216-151	
c	5	102.6	6.2	1205	9	CL143963	CL143963	ISB1-1230	
c	6	102.4	6.2	1162	9	CL077122	CL077122	CH216-143	
c	7	102.2	6.2	1493	9	CL078589	CL078589	CH216-151	
	8	102	6.2	1805	6	CL080711	CL080711	CH216-159	
c	9	101.4	6.1	1626	6	CF238805	CF238805	AGENCOURT	
c	10	101.2	6.1	1542	9	AG386981	AG386981	Mus muscu	
c	11	99.8	6.0	1599	9	CL083840	CL083840	ISB1-2H14	
c	12	99	6.0	1594	9	CL038406	CL038406	CH216-46A	
	13	98.6	6.0	1981	9	CL082000	CL082000	CH216-165	
	14	98.4	5.9	1344	9	CG748432	CG748432	P042-2-C1	
	15	98.4	5.9	1632	9	CL082569	CL082569	CH216-167	
	16	98	5.9	1297	9	CG758143	CG758143	P053-3-B1	
c	17	97.8	5.9	1784	9	CL081992	CL081992	CH216-165	
c	18	97.6	5.9	1101	9	CNS00EVL	CNS00EVL	Drosophil	
c	19	97.6	5.9	1206	8	BZ695529	BZ695529	SP__Ba006	
	20	97.6	5.9	1566	9	CG757757	CG757757	P053-1-D0	
	21	97.2	5.9	1139	8	AQ897537	AQ897537	HS_3153_A	
	22	97	5.9	1638	9	AG347262	AG347262	Mus muscu	
c	23	96.8	5.8	1101	9	CNS0039G	CNS0039G	Drosophil	
c	24	96.6	5.8	1811	9	CG753732	CG753732	P048-4-G0	
c	25	96.4	5.8	1310	9	AG370846	AG370846	Mus muscu	
c	26	96.2	5.8	1512	9	CL082685	CL082685	CH216-169	
c	27	96	5.8	1392	9	CG757503	CG757503	P052-4-C0	
c	28	95.8	5.8	1885	2	BE420745	BE420745	HWM002.B0	
c	29	95.4	5.8	1896	3	CR722884	CR722884	Tetraodon	
	30	95	5.7	1324	9	AG376784	AG376784	Mus muscu	
c	31	94.8	5.7	994	9	CNS04NOJ	AL298972	Tetraodon	
	32	94.6	5.7	1251	9	AG332167	AG332167	Mus muscu	
	33	94.6	5.7	1433	9	AG745119	CG745119	P037-4-G0	
c	34	94.2	5.7	1042	9	CL461494	CL461494	SAIL_1148	
c	35	93.8	5.7	1135	9	CNS033GQ	AL226115	Tetraodon	
c	36	93.8	5.7	1407	9	AJ592026	AJ592026	Arabidops	
c	37	93.4	5.6	938	9	CL509354	CL509354	SAIL_811	
	38	93.4	5.6	1074	8	BZ696936	BZ696936	SP__Ba009	
	39	93.4	5.6	1289	9	AG347097	AG347097	Mus muscu	
	40	93.4	5.6	1348	9	CG749499	CG749499	P043-4-A0	
c	41	93.4	5.6	1489	9	CL078612	CL078612	CH216-151	
c	42	93.2	5.6	1346	9	AG382256	AG382256	Mus muscu	
	43	93.2	5.6	1355	9	AG346348	AG346348	Mus muscu	
	44	93.2	5.6	1539	9	AG340947	AG340947	Mus muscu	
c	45	93.2	5.6	2142	3	CR730230	CR730230	Tetraodon	

```
                                ALIGNMENTS

RESULT 1
AG349719
LOCUS       AG349719           1254 bp    DNA     linear   GSS 02-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-146B10.T7, genomic survey
            sequence.
ACCESSION   AG349719
VERSION     AG349719.1  GI:47923029
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1254)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institute, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : T7
            LIBRARY
            Vector      : pBACe3.6
            R.Site 1    : EcoRI
            R.Site 2    : EcoRI
FEATURES             Location/Qualifiers
     source          1..1254
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-146B10.T7"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
```

ORIGIN

Query Match   6.5%;   Score 107.8;   DB 9;   Length 1254;
Best Local Similarity   47.4%;
Matches 390;   Conservative   0;   Mismatches 427;   Indels   6;   Gaps   2;

```
Qy   279 TTTTATCGTCGAGACATTGGTTACAGGGCTGTGGAACGAGAACAAGCAAAATAGTTAGAA 338
             |||| ||  ||| || ||||||||||| |||||||||  ||| ||||
Db   369 TTTTTTTTATTAAGATGATTATTATTTTTTTTTTTTTAGAAGAAAAAGGATATAATAATATA 428

Qy   339 ATCTTGTTGAATTAATTAGATAAAAAATATGTACAGGGTAGAATTGTATTTTCCTAGG 398
             |||| |||   ||| |||||||||||||| |||||   ||||| ||||||
Db   429 AAATTAATTTTATTGATTTTGTTTTGTGTATATATATGATTTTTTTTTATATT 488

Qy   399 ATTTAGGATTTGTTAGGGCAACGTTACGA-TTGCTCTGACAATAAATTAGAATTATTA 457
             ||||       |||| || |||||||| | ||| || ||||||| |||||||||
Db   489 ATTTGTTTATTTTTATTTGTTTTTTATATTTTTAATTAATGTTTTTTTTT 548

Qy   458 TTTTTGTTACTTTATGAGGTTATATCAACTTGTCATCGTAGTATTATA 517
             |||||| ||| ||||||||  || |||||||||||| ||||||||
Db   549 TTTTTTTTATATTTTTTTATTTTTATTTTTTTTAATTTTTTATTTATTTTATATT 608

Qy   518 TTCTGCAATTCTTCTTTCATTAGAAGTTATTTATTTTATAGAAAAATTATTTCTAGTAT 577
             |||||||||||||||||||||| |||||||| || ||||||||  ||||||
Db   609 TTTTATTATTTATTTTTTTTTTTTATTTTTTATATTTTTTATTATTTTTTTTT 668

Qy   578 TGAGAGATACACTGAAAGTTATTGTCGTCTGTTTGTCGTGTTGTTGTTTTTC 637
             ||||||||||||||| ||| |||||| ||| |||||||||||||| ||
Db   669 TTATTAATTTTTTTTTTTTTTATTTATTTTTTTATTATATTTTATTA 728

Qy   638 AAAATATAGAATTACAAGAATTAGTTGGGCCTTTATTCGCAATAAGTATTGTGTTGAA 697
             ||||||||||    ||||||  |||| ||||||||||||||| ||||||||||||
Db   729 TTATTTTTAATTTATTTATTTTTATTTTATTTTTTTTTTTTTATTTTT 788

Qy   698 TAATGTACATTATGCAGTATACCAATCTGATTGACCCTGTTAATTACTCACTTGCATT 757
             |||| || || || |||||  ||| ||||| |||||| ||| |||||||||
Db   789 TTAATTTTTTATTTTTTTTAATTTTATTTATTTTAATTTTATTTTTTTTT 848

Qy   758 TAAAGAAATTAATGAGATAACCAAATGCTGGCTTAACAATGCTTAACAATGAGATAAAATTCATATATATCC 817
             ||||| | |||| || ||||  ||| || ||||| |||||||| |||||
Db   849 TAATTATTATTTTTATATTTTTTTGTGTGTTTTTTTTTAATTATATTTTTTTATTTTA 908

Qy   818 ATT-----GTTATTTGGTTATTTGAAGTTGCTGTGTGTTTTTAAGTTTAAGTTCATAAAA 872
             |||     ||||||||| ||||| ||||||| ||  |||||||||  ||||||
Db   909 TTTTAATTTTTGTATTTTTGTATTTATTTTTTGTTATATTTTTTTTT 968

Qy   873 AGAAAAGTATATAAAACTTCTTGGATTTTTGACTTTTTTATTTTTTTATGCTGTGATGATGTAT 932
             ||||||||||||||||||  |||| |||| ||||| | ||| ||||||||||
Db   969 AATAATATTTATTGATTTTTTTTTTGTTTTTTATATATTTAATTTTTTTTTAT 1028

Qy   933 GTTTTTTTGTTCGAGCGTATACACAAAATCCCATGAGCGTGTTTTTATTTCACCTAAACACTGTT 992
             ||||||||||||||   | |||||  ||||| ||| |||||| |||||
Db   1029 TATATTTTTATTATTATTATTTATTATTATTATATATATTTTA 1088

Qy   993 TATTCTCCGATTAAAATCCAATTATTATTTTATCGTTGGGTTATTTATAGGACGAATTGTTCCT 1052
             |||    | ||||||||||| |||||||||| ||||||||| ||||||
Db   1089 TATTTATTTTTATTTTTTTTTTATTTATTTTTAATTTTTATTATTTTTATAGTAGTATA 1148

Qy   1053 TATGAGATATTTTCTTTATCCTAAATATTCCCTCTTTATCATCATAAAT 1095
             ||||||   ||||||| |||||||| ||||| ||||||||||||||||
Db   1149 TATTATTTTTTTTATTATTTTTATTAATTTTTATAATTTTATATTTAAT 1191
```

RESULT 2

LOCUS   CL081966   2157 bp   DNA   linear   GSS 31-DEC-2003
DEFINITION   CH216-165D13_Sp5.1 CH216 Xenopus tropicalis genomic clone
      CH216-165D13, genomic survey sequence.
ACCESSION   CL081966
VERSION   CL081966.1   GI:40537879
KEYWORDS   GSS.
SOURCE   Xenopus tropicalis (western clawed frog)
  ORGANISM   Xenopus tropicalis

            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.

REFERENCE   1 (bases 1 to 2157)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE   A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp5   atctgccgtttcgatcct
            Class: BAC ends
            High quality sequence start: 341
            High quality sequence stop: 412.
FEATURES   Location/Qualifiers
     source   1..2157
            /organism="Xenopus tropicalis"
            /mol_type="genomic DNA"
            /strain="Nigerian frog"
            /db_xref="taxon:8364"
            /clone="CH216-165D13"
            /sex="male"
            /cell_line="Stock 248 F7A2, inbred N7"
            /clone_lib="CH216"
            /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
            BAC library"

ORIGIN

Query Match   6.3%;   Score 103.8;   DB 9;   Length 2157;
Best Local Similarity   47.8%;   Pred. No. 3e-09;
Matches 427;   Conservative   0;   Mismatches 458;   Indels   9;   Gaps   4;

```
Qy   350 TTTAATTAGATAAAAAATATTGTACAGGGTAGAATTGTATTTTCCTAGGATTTAGGATTT 409
             ||||| ||||||||||| |||||| |||||||||||||||| ||||||
Db   1148 TTTTATTTATTTATTGTTTTTTTTTTATTTATTTATTTTTTTTATTTT 1207

Qy   410 TGTTAGGGCAACGTTTACGATTGCTCTGACAATAAATTAGAATTATTTTTTGTTACTT 469
             |||||||||   || || |||| || ||||||||| |||| ||||| |||||||
Db   1208 TATTTTATTATTTTTTTTTTTTTGTTATATATTTTTTTTTGTATATT 1267

Qy   470 TATGAGGTTATATCAACTTATGCGACA-ATTGTCATCGTAGTATTATATCTGCAATTC 528
             ||||| ||| || ||||| ||||||| |||||||||| ||||||| ||||||||
Db   1268 TTTTATGTTATTTTATTTTTTTTTTTATTTTATTTTTTTTTTTTTATAT 1327

Qy   529 TTCTTTCATTAGAAGTTATTTATTATAGAAAAATTATTAATATTCTAGTAGTGAGAGATACA 588
             ||||| |||||||||||| ||||||| ||||||| ||||||||  ||||| ||||||||
Db   1328 TATATATTTTTTTGTTTTTTTTAATAATTTTTATTTTTTTTTTTTTTTTTT 1387

Qy   589 CTGAAAGTTATTGTCGTGTTTGTTGTTTTTGTTTATTTTGTTTTTCAAAAATAGAA 648
             |||||| ||||| || ||||||||||||| || |||| ||||| ||||||||||||
Db   1388 TTTATATTTTTGTTTTGTTGTATATTTTTTAATTTATTTTTTTTATATATATTA 1447

Qy   649 TTACAAGAATTAGTTGGGCCTTTATTGTTGCAATAAGTAGTATTTGTGTAATAATGTACATT 708
             ||||||||||||| ||||||||||||||||||| ||||||||| |||||||||||||
Db   1448 TTTTTATTTTATTTATTTTATATATGTTTTGATTTATTTTTTTTTTTT 1507

Qy   709 ATGAGATATACCAATCTGATTGGACCTGTTAATTACTCACTTGCATTTAAAGAAAATTA 768
             ||||||||||||||||||  ||||||| ||||||||||||||||| |||||||||
Db   1508 TTAATTTTTTTTTTATTTTTTTTTTATTTTGTTTTTGTCTTTTATTTTTTTTTTT 1567

Qy   769 ATGAGGATAACAAATGCTGGCTTAACAATGAGATAAATTCATAATATCATTGTTATTTG 828
             ||||  ||||||   |||||||| |||||||||||||||||||||||||||||||||
Db   1568 TTTTTTATTTATTTTGTTTCATATTTTTTATATTTTTTATTTTTTGTATTTTTTTTTA 1627

Qy   829 GTTTATTTGAAGTTGCTGTGTTTTTTTGTATTTTAAGTTTAAGTTTCATAAAAGAAAAAGTATATAAAC 888
             ||||||| ||  ||| ||||||||||| ||| | |||| || |||||||||||||||||||
Db   1628 TTTTTATTTTTTGTTTTTTTGTATTATTTATTTTTTTTTTTTTTTATTTTTATTTATATTT 1687

Qy   889 TTTCTTG---GATTTTGACTTATTTTATGCTGTGATGATGATGTGTTTTTTGTTCGA 944
             |||| ||   |||||| |||| || || ||||||||| ||||||| |||
```

```
Db   382  AATTGTATTTCCAGGATTAGGATTTGTTGTAGGGCAACGTTACGATTGCTCTGACAA       441
Qy  1688  TTTTTGTATTATTATTATTTTTGATTTATTATTTTATTATTTTTATATATATTTATT      1747

Db  1179  TATTAATATATTTTTATAAATATATTAAATTTTTATTTTTATATATAATTATTA         1120
Qy   945  GCGTATACAACAAAATCCCATGAGCGGTTTATTCACCTAACACTGTTATTCGATTA       1004

Db   442  TAAATTAGAATTATTATTTTGTTACTTTATGAGGTTATATCAACTTATGCGACAATTG      501
Qy  1748  TTTTTTGTTTAATATGTTAATTTTGTTTATTTGTGTTTATTTTGTTTTTATTT          1807

Db  1119  TAAATTTATATTTTAATTTAAATATTTTAAAATATTATAATTAATTTTAATATA         1060
Qy  1005  AAATCCAATTATTATCGTTGGGTTATTTATAGGACGAATTGTTCCTATGAGATATT      1064

Db   502  TCATCGTAGTATTATATTCTGCAATTCTTCTTCATTAGAAGTTATTATATAGAAAATTAT   561
Qy  1808  TATTATATATTATTTA--ATTTATTTGTTTATATCTATTCTATTTGTGTATGGTGTTTTATTT  1865

Db  1059  TTATTATTAATAGAATATTTTTTATTTATATTTTTTTTTTATAATAAAT              1000
Qy  1065  TCTTTATCTAATATTCCCTTATCATAAATCTAAGCCTATGAAATCGGGCTCTCCGTAA    1124

Db   562  TTAATATTCTAGTATTGAGAGATACACTGAAAGTTATTGTCGTGTTGTTGTTT          621
Qy  1866  TTTTTTTTTTTATTTGTAATTTGTAATTATTAATTATTTTTTTATTTTTATTTTTATTTT  1923

Db   999  TTATTTATTATTTTTATTTATATAAATAATAATTTAAATAATTTTTATAAAATATT        940
Qy  1125  AATTMARGAATATAATTTTAATTAATGGGGGAAAGTGCGACCTCAAGTCATTTAGTGCT   1184

Db   622  GTTTGTTGTTTTTCAAAAATAGAATTACAAGAAATATATTAGTAGTGGCGCTTATTTGCAA  681
Qy  1924  TGTTTTTGTTTTTATTTAGTTATTTTTTGTTATTTTTATTTATTTATTATT            1983

Db   939  TTATTTTTTAATTTAATTAAATAATTTAAATTATAATTTATGTTAATTTTATAA         880
Qy  1185  TTTGGTTACGGGAGAAAAACATCTCCCTTTTAGATAGCTAAAATAATAAATCA         1238

Db   682  TAAGTATGTTGTTAATAATGTACATTATGCAGTATCGTCAGTACCAATCTTGGACCTGTTA  741
Qy  1984  TTTTTATGATTATTATTAATTATTCTTCTGGTGATATATATTTTTTTTTTTTNATTA     2037

Db   879  -------TTATATAATTTTTTAT                                        828
Qy   742  ATTACTCACTTGCATTTAAAGAAATTAATGAGATAACAAATCTGGCTTAACAATGATAG    801

Db   827  ATATTTAATTTTAATTTATTTATATATATATTTAATTTATTAAAATTAT              768
Qy   802  ATAAATTCATATATCCATTGTATTTGGTTTATTTGTGTTGCTCGTGTTTTAAGTTAA      861

Db   767  ATATATTATTAATTATAATTTTTTAAAAATATTTTATATAAAAATATTAAATTAA        708
Qy   862  GTTTCATAAAAGAAAAGTATTATAAAACTTTCTTGGATTTTGACTTATTTTTATGCTG     921

Db   707  AATTAATTTTATTAAAATTTTTATTTAATTAATATTTTTTATAATATTTAATTAATTTTA   648
Qy   922  TGATGATGTATGTTTTGTTCGAGCGTATCACAACAAAATCCCATGAGCGTTTATTTCAC    981

Db   647  TTATTTTAATTTTTTATATTTATTTTTTAAAATTTATTTTTAAATTTATTTTATTTTTAAA  588
Qy   982  CTAACACTGTTATTCTCGATTAAAAATCCAATTATTCGTTGGGTTATTTATTATAGGAC   1041

Db   587  TAATTTTATATATATAAATAATTTTTATTTTTAAATTTTATAATTATTTTATAAAATATAGT 528
Qy  1042  GAATTGTTCCTTATG---AGATATTTTCTTTATCTAATATTCCTCTTATCATCATAAATCT- 1097

Db   527  TTAATTTTTTATTTATTAATAATTTTTATAATTTTTATAATTATAATAATAAAATA        468
Qy  1098  ------AAGCCTATGAAATCGGGCTCTCCGVAAAAATTMARGAATATAATTTAATTAATGG 1152

Db   467  TAATAAATATATTTAATTTTTTATTAAAAATAATATTATTATATTTTTTATTA          408
Qy  1153  GGGAAGTGCGGACCTCAAGTCATTTTAGTGCTTTTGGTTACGGGAGAAAAAACATCTCCTT 1212

Db   407  TTATAATTATATATTTTTATTATTATTATAATAAAATATTAATAATAATAATTATA       348
Qy  1213  TTTTTAGATAGCTTAAAATAAAATAAAATCAGGAGCGCTCTTGTTGTGTAAAACTTATTCAGGAGGAA 1272

Db   347  ATTTTATATATTAATTTATTAATTTTTGAAAATTTTGAAATTTATTTATTTATTTATTTAT  288
Qy  1273  AGCTAACAGCAATTTCTTTTACCAATGTTTTTAATGCAATYCCTTAYCCAAATGGAATAC  1332

Db   287  AATTTTATTTATTGTTTTATTAGTTTATTTATATTTTTATATTTTTTTAAATTTATAATTAA 228
Qy  1333  AACAGATAGCTAAGGAGGATACGAATTTATTTAATTTA                         1370

Db   227  TTTAAATAATAATATTTTATGTATTTATTTTTTTTTTTA                         190
```

RESULT 3
CC253231/c
LOCUS   CC253231   1277 bp   DNA   linear   GSS 13-MAY-2003
DEFINITION CH261-180N11_RM1.1 CH261 Gallus gallus genomic clone CH261-180N11,
            genomic survey sequence.
ACCESSION   CC253231
VERSION   CC253231.1 GI:30589981
KEYWORDS   GSS.
SOURCE   Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1 (bases 1 to 1277)
AUTHORS   Kremitzki,C., Higginbothan,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE   Gallus gallus BAC End Reads
JOURNAL   Unpublished (2003)
COMMENT   Contact: Richard K. Wilson
            Washington University School of Medicine
            Genome Sequencing Center
            Email: submissions@watson.wustl.edu
            Insert Length: 182000   Std Error: 0.00
            Seq primer: RM1 TACGACTCACTATAGGGAGA
            Class: BAC ends
            High quality sequence start: 24
            High quality sequence stop: 85.
FEATURES   Location/Qualifiers
     source   1..1277
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="CH261-180N11"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="CH261"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CH261 Female Chicken library - for library and clone
                     ordering information: http://www.chori.org/bacpac"
ORIGIN

Query Match   6.2%; Score 103.2; DB 8; Length 1277;
Best Local Similarity   45.7%; Pred. No. 3.9e-09;
Matches   483; Conservative   2; Mismatches   556; Indels   17; Gaps   3;

```
Qy   322  AAGCAAAATAGTTAGAAATCTTCGTTGGAATTTAATTAGATAAAAAATATTGTACAGGGTAG  381
Db  1239  AATATAAATTATATTATTATTATTAATAATAATTATAATTTTTATATTTTTATAATAATAT  1180
```

RESULT 4
CL078538/c
LOCUS   CL078538   1536 bp   DNA   linear   GSS 31-DEC-2003
DEFINITION CH216-151C11_Sp5.1 CH216 Xenopus tropicalis genomic clone

```
ACCESSION   CL078538
VERSION     CL078538.1  GI:40534451
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1 (bases 1 to 1536)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Washington University School of Medicine
            Email: submission@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp5 atctgccgttcgatcct
            Class: BAC ends
            High quality sequence start: 1061
            High quality sequence stop: 1145.
FEATURES             Location/Qualifiers
     source          1..1536
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-151C11"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

Query Match          6.2%;  Score 102.8; DB 9; Length 1536;
Best Local Similarity 46.2%;  Pred. No. 4.6e-09;
Matches 346;  Conservative 0;  Mismatches 402;  Indels 1;  Gaps 1;
```

```
RESULT 5
CL143963/c
LOCUS       CL143963            1205 bp    DNA     linear   GSS 05-JAN-2004
DEFINITION  ISB1-12302_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-12302,
            genomic survey sequence.
ACCESSION   CL143963
VERSION     CL143963.1  GI:40637598
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1 (bases 1 to 1205)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Washington University School of Medicine
            Genome Sequencing Center
            Email: submission@watson.wustl.edu
            Insert Length: 75000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 600
            High quality sequence stop: 664.
FEATURES             Location/Qualifiers
     source          1..1205
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8364"
                     /clone="ISB1-12302"
                     /clone_lib="ISB1"
                     /note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                     Library Segment 1"
ORIGIN

Query Match          6.2%;  Score 102.6; DB 9; Length 1205;
Best Local Similarity 44.0%;  Pred. No. 5e-09;
Matches 349;  Conservative 2;  Mismatches 440;  Indels 2;  Gaps 1;
```

(Sequence alignment Qy/Db blocks with position numbers 340–819 / 1478–1000 for the first result and 340–608 / 399–549 for RESULT 5 follow; the aligned nucleotide traces are shown with alignment bars between Qy and Db rows.)

ACCESSION row header (top): CH216-151C11, genomic survey sequence.

High quality sequence stop: 808.

FEATURES   Location/Qualifiers
source   1..1162
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-143E5"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"

ORIGIN

Query Match   6.2%;   Score 102.4; DB 9;   Length 1162;
Best Local Similarity 46.5%;   Pred. No. 5.5e-09;
Matches 328; Conservative   0; Mismatches 377; Indels   0; Gaps   0;

Qy   384 TTGTATTTCCTAGGATTAGGATTTGTAGGGCAACGTTACGATTGCTCTGACAATA 443
          ||||||||||| ||||| | ||||| | |||||||| |||||||||||||
Db   1120 TTCTAATTTATTATTTTTTAATTTTATTATGGTTATCAACAACTTATCTTTATCATA 1061

Qy   444 AATTAGAATTATATTTTGTTACTTATCTTATGAGGTTATATCAACTTATGCGACAATTGTC 503
          |||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db   1060 TTTATTTTATTTTATTTTTTTTTTTTTTATTTTTTTT 1001

Qy   504 ATCGTAGTATTATATTCTGCCAATCCTCTTCTTCATTAGAAGTTATTATAGAAAAATTATTT 563
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1000 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTT 941

Qy   564 AATATTTCAGTAGTATTGAGAGATACACTGAAAGTTATTGTCTGTTGTTGTTGT 623
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   940 TATTATTTTTTTAATTTATTTTTATTTTTTTATTTTTTTTTTTTTTTTTTTTT 881

Qy   624 TTGTTGTTTTTCAAAATATAGAATTACAAGAATATTAGTTGGCGCTTATTGCAATA 683
          |||||||||||||||||||||||||||
Db   880 TATTTTTTTATTTTTTTTTTTTTTTTTTATAATTTTA 821

Qy   684 AGTATTGTTGTTAATAATGTACATTATGCAGTACGAGTATACCAATCTGGATTGGACCTGTTAAT 743
          ||||||||||||||||||||||||||||||||||||||||||
Db   820 TATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 761

Qy   744 TACTCACTTGCATTTAAAGAAATTAATGAGATAACAAAATGCTAGCTTAACAATGATAGAT 803
          |||||||||||||||||||||||||||||||||||||||||||
Db   760 TTTTCTTTTTTTTTTTTTTTTTATATNATTTTTTTTTTTTTTTTTTATTTTT 701

Qy   804 AAATTCATATATCCATTGTTATTTGGTTTATTTGAAGTTGCTGTGTTTTAAGTTTAAGT 863
          |||||||||||||||||||||||
Db   700 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTATTTTTAT 641

Qy   864 TTCATAAAAGAAAAGTATATAAACTTCCTGGATTTTGACTTTATTTTTATTTTTTATGCTGTG 923
          |||||||| |||| || ||||||||||||||||||||||||||||||||
Db   640 TTTTATTTTATTTTTTTTTTTTTATTTTTTATTTTTATTTTTTTTTTTTTTTTTTA 581

Qy   924 ATGATGTATGTTTTGTTCGAGCGTATACAACAACAAAATCCATGAGCGTTTATTCACCT 983
          ||||||||| |||||||||||||||||||||||||||||||
Db   580 TTTTTTTTTT|||||||TTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTT 521

Qy   984 AACACTGTTTATTCCGATTAAAATCCAATTATTATCGTTGGGTTATTTGTTGGGTTATTTTATAGGACGA 1043
          ||||| ||||||||||||||||||||||||||||||
Db   520 TTTTTTTTTTTTTTTTTTTTTATTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTT 461

Qy   1044 ATTGTTCCTTATGAGATATTTCTTTATCTAATATTCCTCTTAT 1088
          |||||||||| |||| || | |||||
Db   460 TTTTTTTTTTTTTTTTTTTATTTTTTATTTATTTTTTTTTTTTTTTTT 416

RESULT 7
CL078589/c
LOCUS   CL078589   1493 bp   DNA   linear   GSS 31-DEC-2003
DEFINITION CH216-151K5_Sp5.1 CH216 Xenopus tropicalis genomic clone
            CH216-151K5, genomic survey sequence.

───

Db   1082 TATATTTTTATTTTATTATTAATATTTTTATTATTTTAATTTTTTTTTTTTATTT 1023

Qy   609 TGTTTGTTGTTGTTGTTGTTTCAAAAATAGAATTCAAGAGATATTAGTTGGC 668
          | ||||| || ||| || || ||||||| | ||||||
Db   1022 TTTTTTTTTTTTTATTTTNTTTTATTTTTTTTTTT 963

Qy   669 GCTTATTGCAATAAGTAGTTGTGTAATAATGTACATTATGCAGTATACCAATCTTGG 728
          | ||||||||||||||||||||||||||||||||
Db   962 TTTTTTTTTTTTTTTATTTTNTNTTTTTTTTTTTTATTTTATTTTTTNTTTTTATTT 903

Qy   729 ATTGGACCTGTTAATTACTACTCACTTGCATTTAAAGAAATTAATGAGATAACAAATGCTGGC 788
          ||||||||||||||||||||||||||||||||||||
Db   902 TTTTTTTTTTTNTTTT--TTTTTTNNTTTTTTTTTTTTTTTTTTTTTTATAATATTATA 845

Qy   789 TTAACAATGATAGATAAAATTCATATATCCATTGTTATTTGGTTTATTTGAAGTTGCTGTG 848
          |||||||||||||||||||||||||||||| |||||||||||||||
Db   844 TTTATATTTTNTNTTTTTTTTTATTATTTTTTTTTTTTTTTTTTTNTTTTT 785

Qy   849 TTTTTAAGTTAAGTTCATAAAAAGAAAAGTATATAAACTTCTTGGATTTTTGACTTT 908
          ||||| |||||||||||||||||||||||||||||||
Db   784 TTTTTTTTTNTTTTATTTTTTTTTTTTTTTTTTTNTNTTTTTTTTTT 725

Qy   909 ATTTTTATGCTGTGATGATGATGTATGTTTTGTTCGAGCGTATACAACACAAAATCCCATAG 968
          |||||||||||||||||||||||||||||||||||
Db   724 TTTTTTTTTTTTTTTTTTTTTTTTNTTTNTNTTTTTTTTTTTTTTTTTTTTTT 665

Qy   969 CGTTTATTCACCTAACACTGTTATTCTCGATTAAAATCCAATTATTATCGTTGGGT 1028
          |||||||||||||||||||||||||||||||||
Db   664 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTAT 605

Qy   1029 TATTTTATAGGACGAATTGTTCCTTATGAGATATTTTCTTATCTAATATTCCTCTTTAT 1088
          |||||||||||||||||||||||||||||||||
Db   604 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 545

Qy   1089 CATAAAATCTAAGCCTATGAAATCGGGCTCCTCCGVAAAAATTMARGAATATAATTTTAATTA 1148
          ||||||||| | |||| |||||||| ||||
Db   544 TTTTTTTTTTTTTTTTNTTNTNNTNNTTTTTTTTTAAAAAAATAAAATANTTNNTNNNNT 485

Qy   1149 ATGGGGAAAGTGCGACCTCAAGTCATTTAGTGCTTTTGGTTACGGGAGAAAACATCT 1208
          ||||||||||||||||||||||||||||
Db   484 NTTNANNNNNNNAAAAAAAANTTTTTTTTTTTNAAAAAAAANNNNTNNNTNT 425

Qy   1209 CCTTTTTAGATA 1221
          ||||| |
Db   424 NNNTTTTTAAAA 412

RESULT 6
CL077122/c
LOCUS   CL077122   1162 bp   DNA   linear   GSS 31-DEC-2003
DEFINITION CH216-143E5_Sp6.1 CH216 Xenopus tropicalis genomic clone
            CH216-143E5, genomic survey sequence.
ACCESSION   CL077122
VERSION   CL077122.1   GI:40533035
KEYWORDS   GSS.
SOURCE   Xenopus tropicalis (western clawed frog)
  ORGANISM Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1 (bases 1 to 1162)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE   A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            High quality sequence start: 757

```
ACCESSION   CL078589
VERSION     CL078589.1  GI:40534502
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1493)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000    Std Error: 0.00
            Seq primer: Sp6 ctcaaggcgtttcgatcct
            Class: BAC ends
            High quality sequence start: 1009
            High quality sequence stop: 1070.
FEATURES             Location/Qualifiers
     source          1..1493
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-151K5"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

Query Match          6.2%;   Score 102.2;  DB 9;   Length 1493;
Best Local Similarity 46.6%;  Pred. No. 6e-09;
Matches  327;  Conservative  0;  Mismatches 372;  Indels  3;  Gaps  1;

Qy    387 TATTTTCCTAGGATTTAGGATTTTGTTAGGGCAACGTTTACGATTGCTCTGACAATAAAT 446
             |||||||| |||||||    ||| || | ||||||||||    |||  | |||||||
Db   1142 TTTTTTAATTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAT 1083

Qy    447 TAGAAATTATATTTTTGTTACTTTATGAGGTTATATCAACTTATCGCGACAATTGTCATC 506
             ||||||  |  || || ||||||||||   ||  |||
Db   1082 TTTTTTTATTATTTTTTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAT 1023

Qy    507 GTAGTATTATTCTGCAATTCTTCTTCATTAGAAGTTATTATAGAAAATTATTTAAT 566
             ||||| |      |||||||||||||  |||||
Db   1022 TTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTAATAATTTTTT 963

Qy    567 ATTTCTAGTAGTTGAGAGATACACTGAAAGTTATTGTCTGTTTGTTGTTGTTGTTTG 626
             |||||||||||| || |||||||||||| |     ||| |||  |||||||||||
Db    962 TTTTTTTAATAATTTTTTTTTTTTTTTTTTTATTTTTTTTTTTATTTTTTTTTTTT 903

Qy    627 TTTGTTTTTTCAAAATATAGAATTACAGAAGATATTAGTTGGCGCCTTATTGCAATAAGT 686
             |||||     |||||
Db    902 TTTTTTTTTTTTTTTTTNTTTTTTTANTTTTTTT---TTTTTTTTTTTTT 846

Qy    687 ATTGTTGTTAATAAATGTACATTATGCAGTATACCAATCTTGGATTGGACCTGTTAATTAC 746
             ||||| ||  | ||
Db    845 TTTTTTTTATTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTT 786

Qy    747 TCACTTGCATTTAAAGAAATTAAATGAGAAGAAACAAATGCTGGCTTAACAATGACATGAA 806
             ||||||||
Db    785 TTTATTTAAATAAAATATAATTTTTTTTTTTTTTNTTTTTTTTTTTTTTATATA 726

Qy    807 TTCATATATCCATTGTTATTTGGTTATTGAAGTTATTGAAGTTGCTGTTTTAAGTTAAGTTTC 866
             |  |||||||||  |||
Db    725 TTTATTTTTTTTTTTTTTTTT ||||||||||||TTTTTTTTTTTTTTTTNTTT 666

Qy    867 ATAAAAGAAAAGTATATAAACTTTCTTGGATTTTGACTTATTTTTGACTTATTTTTATGCTGTGATG 926
```

```
Db    665 TAATAAAAAAAAAAANTTTTTTTTTTTNTTTTTTTTTTTNTTTTT 606
             |||| | ||||||||||         |        ||    ||
Qy    927 ATGTATGTTTTGTTCGAGCGTATCACAACAAATCCATGAGCGTTTATTCACCTAAC 986
             |||||||
Db    605 TTTTTTTTTTTTTTTTTTTTTTTNTTTTTNTTTTTTTTTTTT 546
             |||   |||
Qy    987 ACTGTTTATTCTCGATTAAAATCCAATTATTATCGTTGGGTTATTTATAGGACGAATT 1046
             |||||||
Db    545 NTTTTTTTTTTTTTTTTTTTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTN 486
             ||
Qy   1047 GTTCCTTATGAGAATATTTCTTATCTAATATTCCTCTTAT 1088
             |||||
Db    485 NTTTTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNT 444
```

```
RESULT 8
CL080711
ACCESSION   CL080711
VERSION     CL080711.1  GI:40536624
KEYWORDS    GSS.
LOCUS       CH216-159I8_RM4.1 CH216 Xenopus tropicalis genomic clone
            CH216-159I8, genomic survey sequence.        1805 bp    DNA    linear    GSS 31-DEC-2003
DEFINITION  CH216-159I8_RM4.1 CH216 Xenopus tropicalis genomic survey sequence.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1805)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000    Std Error: 0.00
            Seq primer: RM4 ctcaagggcatcggtcgagc
            Class: BAC ends
            High quality sequence start: 690
            High quality sequence stop: 766.
FEATURES             Location/Qualifiers
     source          1..1805
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-159I8"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN

Query Match          6.2%;   Score 102;  DB 9;   Length 1805;
Best Local Similarity 43.3%;  Pred. No. 6.5e-09;
Matches  546;  Conservative  4;  Mismatches 706;  Indels  4;  Gaps  2;

Qy    329 ATAGTTAGAAATCTTGTTGAATTAATTAGATAAAAAAATATTGTACAGGGTAGAATTGTA 388
             ||||||||
Db    466 ATATTATTTATTTATTATTGATTTTTTTTTTTATATTTTTTTATAATTTTTTCNTNTATGT 525

Qy    389 TTTTTCCTAGGATTTAGGATTTTGTTAGGGCAACGTTTACGATTGTGCTGCTGACAATAAATTA 448
             |||||
Db    526 TTTTTTATATTTTTTATATAATATTTATTTTTTTTTTAATATCACCATTACAATAATTTATA 585

Qy    449 GAATTATTATTTTGTTACTTATTATTCACTTTATTATTCACTTTATTTTTTATTTTTATTCGTGATCGT 508
             |
Db    586 CATATTTACATTTATATTATTATTTACTTTATTATTATTATTTTTTTATTTTATTT 645
```

RESULT 9
CF238805
CF238805/c

LOCUS   CF238805   1626 bp   mRNA   linear   EST 05-AUG-2003
DEFINITION   AGENCOURT 15099447 NICHD_XGC_Emb6 Xenopus tropicalis cDNA clone
                IMAGE:6995950 5', mRNA sequence.
ACCESSION   CF238805
VERSION   CF238805.1 GI:33442013
KEYWORDS   EST.
SOURCE   Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
                Xenopodinae; Xenopus; Silurana.
REFERENCE   1 (bases 1 to 1626)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Daniela S. Gerhard, Ph.D.
                Office of Cancer Genomics
                National Cancer Institute / NIH
                Bldg. 31 Rm10A07 Bethesda, MD 20892
                Email: cgapbs-r@mail.nih.gov
                Tissue Procurement: Robert M. Grainger
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Agencourt Bioscience Corporation
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLAM14680 row: g column: 21
                High quality sequence start: 71
                High quality sequence stop: 316.
FEATURES   Location/Qualifiers
     source   1..1626
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:6995950"
                     /tissue_type="neurula"
                     /dev_stage="embryo, stages 14-19"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NICHD_XGC_Emb6"
                     /note="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;
                     Cloned unidirectionally. Primer: Oligo dT. Average insert
                     size 2.1 kb. Constructed by Invitrogen. Note: This is a
                     Xenopus Gene Collection (XGC) library."

ORIGIN

Query Match   6.1%;   Score 101.4; DB 6; Length 1626;
Best Local Similarity 40.2%;   Pred. No. 8.4e-09;
Matches 498; Conservative 3; Mismatches 739; Indels 0; Gaps 0;

Qy   509 AGTATTATATTCTGCCAATCTCTTCCTTCATTAGAGAGTTATTATTTTTTATAGAAAATTATTTAATAT 568
         || || || |||||
Db   646 ATTTTTTTTATTTTTTTTATTTTTTTTATTATATATTTTTGTTGTTGTTGTTGTTGTTGT 705

Qy   569 TTCTAGTATTGAGAGATACACTGAAAGTTATTTGTCTCGTTGTTGTTGTTGTTGTTGTTGT 628
         ||
Db   706 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTT 765

Qy   629 TGTTTTTCAAAATATAGAATTACAAGAGATATAGTTGGCGCTTATTGCAATAAGTAT 688
         || ||||
Db   766 TTTCTTTTTTTTTTTTATTCTATATAATTTTTTTTTTTTTCTATTTTTT 825

Qy   689 TGTGTGTTAATAATGTACATTATGCAGTATACCAATCTTGGATTGGACCTGTTAATTACTC 748
         ||| |||
Db   826 TTTTTATTTTTTTTTTTTCATTATATTTTTTTTCTTTATTTTTTTTTTTTTT 885

Qy   749 ACTTGCATTTAAGAAATTAAGATGAGATAACAAATGCTCGGCCTTAACAATGATAGATAAATT 808
         ||
Db   886 TTATTTTTTGTTCCTAATTTTTTTTTTTTTTTTTTTTATTTTTAGTTATATAATTTATTATTTT 945

Qy   809 CATATATCCATTGTTATTGGTTTATTTGAAGTTGCTGTGTTTTTTTTAAGTTTCAT 868
         || ||
Db   946 TATTTATTTTTTTAATTTTTTCTTTTTTTTTTTTTTTTAATATCTATATTT 1005

Qy   869 AAAAAGAAAAGTATATAAACTTTCTTGGATTTTGACTTTATTTTATTATGCTGTGATGAT 928
         |||||||||||
Db 1006 AATTATTTTATTATAAAATATTTTTCTTTTTTTTTAATTTTTTTAAATTTTTTTATTT 1065

Qy   929 GTATGTTTTTGTTCTGAGCGTATACACAACAAAATCCCATGAGCGCGTTTATTTCACCTAACAC 988
         |||
Db 1066 TTTTTTCATTTTATTAATTATCCAATTATCTCGTTGGGTTATTTTCATTTTTATAGGACGAATTGT 1125

Qy   989 TGTTTAATTCTCGATTAAAATCCAATTAATTTTTTTTTTTTTTTCATTTTATTTTATTAATAATATTGT 1048
         | |||||| |||
Db 1126 TTTTTATTTT--TTTAATATTTTTTTTTTTATTTTTCATTTTATTTTATTAATAATATATTTT 1183

Qy 1049 TCCTTATGAGATATTTTTCTTTATCTATAATTCCTCTTTATCATAAATCTAAGCCTATGAA 1108
         || |
Db 1184 TTTTATTTTTAATCAAATATTTTTTATTATATTTAGTTTACAACACATATATTATTTTTT 1243

Qy 1109 ATCGGGCTCTCCGVAAAATTMARGAACATAATAATTTAATTTATATTTAATGGGGGAAAGTGCGACCTC 1168
         |||
Db 1244 TATATTTTACTATTTTATTTACATAAATTTATTTATATAAATTTTACATAATTTAAATTTTAAGGG 1303

Qy 1169 AAGTCATTTTAGTGTCTTTTGGTTACGGGAGAAAAACATCTCCTTTTTAGATAGCTTAAA 1228
         || |
Db 1304 TCTGNATTTTAATATTTTAATTATTTAATAACATTTATTTACTTTTATTATTTAAAAAATATATCAAT 1363

Qy 1229 ATATAAATCAGGAGCTCTTGTTCGGTAAAAACTTATTCAGGAGGAAAGCTAACAGCAATTC 1288
         ||
Db 1364 TATTATATTATTTCTATTTTTTTTTTTTTTTATATTTCTATATTAAAATTATTTATTTATAA 1423

Qy 1289 TTTACCAATGTTTTTAAATGCA--ATYCCTTAYCCAAATGGAATACAACAGATAGCTAAA 1346
         || | |: || :
Db 1424 TTTTTTTTNAATTTTTTTTTTCTAAATATTTTTTATATAATTTTTTTATATANTTTTATTTTATTA 1483

Qy 1347 GGAGATACGCAATTTATTTAATTTAGCGAAAGAGCAAGGCTTCAGACATATTTTTTATTCA 1406
         || |||
Db 1484 AAAAATGATTTATTTATTTTTTTTTTTTTTGTATTTTACTTT 1543

Qy 1407 GCTCAAGCTAGGGATGATGATGATCGATATGCATCAATTTTTAGGAGGAGCTTGGATTGATGAT 1466
         ||| ||
Db 1544 TATTTTTTAAAATTTTTTTTTTTACATATTATTAATTATATATTTTATTATTA 1603

Qy 1467 ATTCGTTTTCCAGATAATGAAGGGTATTCTTAAGAGAGATTCAATGCCTGATAATAAAATTA 1526
         || |||
Db 1604 AATTTTTTTTATAATTAARTAATTGTTTGTAAATTTTTTATAAATATTATTTATTTTTAAATT 1663

Qy 1527 CTTCCTGCCTTTAAAAAATATTATTTTTTATTAGTAGAATATTAGTGTATCATTTTGTTGTTGTTACATCAT 1586
         || ||
Db 1664 TTACATATTTTTATTTTTTTTTTTTATCATTTTTTTTTTTATTTATTTTTTTATTTATTCTTTTTTTAT 1723

Qy   640 AATATAGAATTACAAGAGATATTAGTTGGCGCTTTATTTGCAATAAGTATTGTTGTTAATA 699

```
ORGANISM  Mus musculus molossinus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE     BAC end Sequences of Library MSMg01
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 1542)
AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
TITLE     Direct Submission
JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
          (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
          Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT   Clones are derived from the mouse BAC library MSMg01. For BAC
          library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
          Tsukuba Institude, Bio Resource Center,
          The Institute of Physical and Chemical Research (RIKEN) 3-1-1
          Koyadai, Tsukuba, 305-0074 Japan
          phone: 81-298-36-9189, fax: 81-298-36-9199
          e-mail: abe@rtc.riken.jp
          PRIMERS
          Sequencing : TJ
          LIBRARY
          Vector     : pBACe3.6
          R.Site 1   : EcoRI.
          R.Site 2   : EcoRI.
FEATURES  Location/Qualifiers
     source  1..1542
             /organism="Mus musculus molossinus"
             /mol_type="genomic DNA"
             /sub_species="molossinus"
             /db_xref="taxon:57486"
             /clone="MSMg01-201G10.TJ"
             /sex="male"
             /tissue_type="mixture of kidney and spleen"
             /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match           6.1%;  Score 101.2;  DB 9;  Length 1542;
Best Local Similarity 44.4%;
Matches 439;  Conservative   3;  Mismatches 544;  Indels  3;  Gaps  2;
```

```
Qy  333 TTAGAAATCTTGTTGAATTAATTAGATAAAAAATATTGTACAGGTAGAGATTGTATTTT 392
            ||||| ||         |||||        |||||||                    
Db  420 TTATAAATTTTTTTTTATATTTTTATAAAATATTTTTTTTTATATATAAAATATA 479

Qy  393 CCTAGGATTTAGGATTTGTTAGGGCAACGTTACCGATTGCTCTGCACAATAAATTAGAAT 452
            | ||||       ||||||||              ||||                  
Db  480 AATAATATATTTTTTTTTTATATATATAAATTTATATATATTTATTTTTTTTATATT 539

Qy  453 TATTATTTTTGTTACTTTATGAGGTTATATCAACTTATGCGGACAATTTGTCATCGTAGTA 512
            ||||||         |||||||||||                               
Db  540 TTTTTATTTTTTTTATATATTTTATATATTTTTAATTTTTAATATANATTAATAATA 599

Qy  513 TTATATTCGCAATCTTCTTCATCAGAAGTTATTATTATTAATAGAAAATTATTTAATATTTCT 572
            |||||         |||||||            ||||||  ||||            
Db  600 TTTAAATATTTTATTTTTATTTTTTTATTTTTTAAAA-TATTATTTATTATTATT 658

Qy  573 AGTATTGAGAGATACACTGAAAGTTATTTGTCGTGTTTTGTTGTGTTTTGTTGGTT 632
            |||||||                 |||||                            
Db  659 TTTATTTATAAATATATATTTTTATTATTATTTATTTTTTTTTATATATATTATT 718

Qy  633 TTTTCAAAATATAGAATTACAAGAGATATTAGTTGGCGCGCTTATTGCAATAAGTATTGTT 692
            ||||                |||||||           ||||                
Db  719 TTTTTTTATTATATATTTATTTTATTTATTTTTTATTTTTNATATTTTTATTT 778

Qy  693 GTTAATAAATGTACATTATGCGATACGAGTATACCAATCTTGGATTGGACCTGTTAATTACTCCACTT 752
            |||||                   |||||           |||||            
Db  779 TTTTTTTATTTTTATTTTTTATTTTTTATTAATTAATTTTTATAATATTATTATT 838

Qy  753 GCATTTAAAGAAAATTAATGAGATAACAAATGCTGGCTTAACAATGATAGATAAAATTCATA 812
```

```
Db 1311 ANTTTTTTTTTTTTTTTTATTTTTTTTTTTTATATTTTTTTTTTTTTA 1252
            ||||            |||| |||||        
Qy  700 ATGTACATTATGCAGTATACCAATCTTGGATTGGACCTGTTAATTACTCACTGCATTTA 759
            |                  ||||                
Db 1251 TTTTTTTTTTTTTTTTTTTTTTATTATTATATATATATATTTTTTTTTTTTT 1192

Qy  760 AAGAAAATTAATGAGATAACAAATGCTGGCTTAACAATGATAGATAAATTCATATATCCAT 819
            ||                |||                
Db 1191 TTTTTTTTTTTTTATTTATTTTTTTTTTTTTTTTTTTTTTTTTT 1132

Qy  820 TGTTATTGGTTTATTTGAAGTGCTGTGTTTTTTAAGTTCATAAAAAGAAAAAG 879
            ||||                |||||                
Db 1131 TTAATTTTTTTTTTTTTTATTTTTTTTTTTATTTTTTTTTTAATTTTTTTTTATTTT 1072

Qy  880 TATATAAACTTTCTTGGATTTTGACTTTATTTTTTTATGCTGTGATGAGATGATGTTTTTG 939
            |||||          |||||||||                
Db 1071 TTTTTTTATTTTTATTTTTTTTTTATTTTTTTTTNATTTTTTTTTC 1012

Qy  940 TTCGAGCGTATACAACACAAAATCCCATGAGCGTTTATTCACCTAACACTGTTTATTCTC 999
            |                   |||                
Db 1011 TTTTTATTTTTTTTTTTATATTATTTTTTTTTTTAYTTTTTTTTTTTTTATTTTTT 952

Qy 1000 GATTAAAAATCCAATTATTATTATCGTTGGGTTATTTATAGGACGAATTGTTCCTATGAGA 1059
            ||||||                |||||                
Db  951 TTTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTAAATATTTTTTTTTTTTTTTTT 892

Qy 1060 TATTTTTCTTTATCTAATATTCCTCTTATCATAAATCTAAGCCTATGAAATCGGGCTCTC 1119
            |||||                |||||                
Db  891 TTTTTATTTTTAANTTTTTTTAATTTTNTNTNATTATTTTTTTTTTTTTTANTNNTTTTT 832

Qy 1120 CGVAAAATTMARGAATATATATTTTAATTAATGGGGAAAGTGCGACCTCAAGTCATTTTA 1179
            |                   |||||                
Db  831 TTTTTAAAAATTTTTTTTTTTTTTTTTTTNNNTTTTTTTTNNTTTTTTTTTT 772

Qy 1180 GTGCTTTGGTTACGGGGAGAAAAACATCTCCCTTTTAGAGTAGCTAAAATATAAATCAG 1239
            ||||||                |||||                
Db  771 TTTTTTTTTTTTTTTTTAATTNNATAAAATTTNTTTTTTTTTTTTTTTTTTTTAC 712

Qy 1240 GAGCTCTGTTGGTAAAACTTATTCAGGAGGAAAGCTAACAGCAATTCTTTACCCAAATGT 1299
            |||||                |||||                
Db  711 TTTTAATTATTATTTTTTTTTTTNANATTNTNTNTTTTTTTTTTTTTTTTTATTTTT 652

Qy 1300 TTTTTAATGCAATYCCTTAYCCAAATGGAATACAACAGATAGCTAAAGGAGATACGAATT 1359
            |||||                |||||                
Db  651 TTTTTTTTTTTTTTTTTTTTTTTANNANANATAATTATATATTTTTTTTTTTTTTT 592

Qy 1360 TATTTAATTAGCGAAAGAGCAAGAGGCTTCAGACATATATTTTTATTCAGCTCCAAGCTAGGG 1419
            |||||                |||||                
Db  591 TTTTTNNNANNNNNAAAAAAAATATTTATAAATTTTTTTTTTTTTTNTNNNNNNN 532

Qy 1420 ATGATATGCATATGATCAATTTTTAGGAGGAGCTTGGATTGATGAGATATTCGTTTTCCAG 1479
            ||||||                |||||                
Db  531 NNNAAAAAAAAATAAAATTTAAAATTTTTTTTTTTTTTNNNNANANAAAAAAAAA 472

Qy 1480 ATAATGAAGGGTATTCTTTAAGAGATTCAATGCCTGATAATAAATAACTTCCTGCTTTTA 1539
            ||||||                |||||                
Db  471 AAAANAAANAAAAAAAAAATTTATNTTTTTTTTTTTTTAANAANANAANAAANAAANAA 412

Qy 1540 AAAATATTAATTTAGATAATGGTTATCATTTTGTTGTTT 1579
            ||||||                |||||                
Db  411 AAAAATTTTTTTATATATATCTCTATATATATATATTTT 372
```

```
RESULT 10
AG386981
LOCUS       AG386981             1542 bp    DNA     linear   GSS 03-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-201G10.TJ, genomic survey
            sequence.
ACCESSION   AG386981
VERSION     AG386981.1  GI:47998186
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
```

Mon Dec 27 09:35:03 2004

us-10-698-235-1.rst

```
                    /clone_lib="ISB1"
                    /note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                     Library Segment 1"

ORIGIN

Query Match          6.0%;    Score 99.8; DB 9;  Length 1599;
Best Local Similarity 46.3%;  Pred. No. 1.7e-08;
Matches 326;  Conservative  0;  Mismatches 378;  Indels  0;  Gaps  0;
```

```
Qy  ATTGTATTTCCTAGGATTAGGATTTGTTAGGGCAACGTTACGATTGCTCTGACAAT 442
    ||| ||||
Db  AFTATTTTATTATTATTATTTATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1337

Qy  AAATTAGAATTATATTTTGTTACTTATGAGGTTATCAACAACTTATGCGACAATTGT 502
       ||                ||| |
Db  TTTTTTTATTTATTTATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTT 1277

Qy  CATCGTAGTATATATTCTGCAATCTTCTTTCATTAGAGAGTTATTATAGAAAATTATT 562
    ||| |||                |||
Db  TTTTTTTATTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTT 1217

Qy  TAATATTTCTAGTATTGAGAGATACACTGAAAGTTATTTGTCTGTTGTTGTTGTTTG 622
    |||||                      |
Db  TTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1157

Qy  TTTGTTGTTTTTCAAAATATAGAATTACAAGAATATTAGTTGGCGCTTATTGCAAT 682
    |||||||||               |||
Db  TTTTTTTTTATTTATTTTNTTATTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTT 1097

Qy  AAGTATTGTGTTAATAATGTACATATGCAGTATACCAATCTTGGATTTGGACCTGTTAA 742
    |||||||||                |||
Db  TTTTTTTTTTTTTTTTTATTTTTTTTATTTTTATTTTTTTTTATTTTT 1037

Qy  TTACTCACTGCATTAAAGAAATTAATGAGATAACAAATGCTGGCTAACAATGATAGA 802
    |||||  |||
Db  TTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 977

Qy  TAAATTCATATATCCATTGTTATTTGGTTATTTGGTTTATTTGAAGTGTCGTGTGTT 862
    |||||| |||
Db  TTTTTATATTTTTTTTTTTTTTTTTTNTTTTTTTTTTTTTTTTTTTTTTTAT 917

Qy  TTTCATAAAAGAGAAAGTATATAACTTTCTTGGATTTTTGACTTTATTTTTATGCTGT 922
    |||       || |||||
Db  TTTTTTTTTTTATTTATTTTTTATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 857

Qy  GATGATGTATGTTTTGTTCGAGCGTATACAACAAAATCCCATGAGCGCGTTTATTTCACC 982
    |||| ||||||
Db  TTTTTTTTTTATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTT 797

Qy  TAACACTGTTTATTCTCGATTAAAATCCAATTATTTATCGTTGGGTTATTTTATAGGACG 1042
    ||| ||||||
Db  TTTTAATTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 737

Qy  AATTGTTCCTTATGAGATATTTTCTTTATCCTAATATTCCTCTTT 1086
    |||||| |||
Db  TTTTTTTTTTATTTTTTTTTTTTTTTATATTTTTTTTTTTTTTTT 693
```

```
Db  ATTTATTTTAATTTTTAAATTTAAAATATATTAATTATTTATTTATAAGTTTATTATT 898
    ||| |||          ||||
Qy  TATCCATTGTTATTTGGTTATTGAAGTGTCGTGTTTTTAAGTTTCATAAAAA 872
    ||||||
Db  TTTTTTAATATTTATTATTATTATTTAATTNTTTTTTTTTATATTATTATTATTA 958
    ||||||
Qy  AGAAAAGTATATAAACTTCTTGGATTTTTGACTTTATTTTTATTTTTATGCTGATGATGTAT 932
    ||||||||
Db  TTAATATATATATATATATATTTTATTTATTTTAAAATATTATTTTTTTTTTATATT 1018
    |||||||
Qy  GTTTTGTTCGAGCGTATACAACAAAATCCCATGAGCCGGTTTATTTTATATATTTTTTTAT 990
       ||||||||
Db  TTTATTTTTTATTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTAT 1078
    ||||||||
Qy  TTATTTTCTCGATTAAAAATCCAATTATTTTCTTATTTANTATTTATATTTAATATATTTT 1050
       |||||||
Db  ATTTATTTATTATAATTTTCTTTAATTTTTATCTCATAAAAATCTAAGCCTATGAAAT 1138
       ||||||
Qy  CTTATGAGATATTTCTTTATTCTCCTTCTTTATCATAAATCTAAGCCTATGAAAT 1110
       ||||||
Db  ATTTAAATTTATATANTATTTTATTTTTAATATNTAATATATAATAATTAT 1198
    ||||||
Qy  CGGGCTCTCCGTAAAATTMARGAATATAATATTTAATTAATAATGGGGAAGTGCGACCTCAA 1170
    ||||| 
Db  ATTATTTTTATNATTTTTTTNTTTATTTATTATTAATAATATTTTAATTA 1258
       |||||||
Qy  GTCATTTTAGTGTCTTTGGTTACGGGAGAAAAACATCTCCCTTTTTAGATAGCTTAAAAT 1230
    ||||||||
Db  TTTATTTTATTTTTTATTTTNTAAATATTTTTTTTATTTATATATATATATTTT 1318
    ||||||
Qy  ATAAATCAGGAGCTCTTCGTTGGTAAAAACTTATTCAGGAGGAAAGCTAACAGCAATTTCTT 1290
    |||||||
Db  AATTNTAATTATTATATATATATTATTTAATTATTTACTATATNTTTTTTATTTTTTAATAT 1378
    ||||||
Qy  TACCAAATGTTTTTAATGCAATYCCTTAY 1319
    ||| |||
Db  TTTNTAATNTATATATATATTATTTATTTATTTTAT 1407
```

RESULT 11
CL083840/c

```
LOCUS       CL083840    1599 bp    DNA    linear   GSS 05-JAN-2004
DEFINITION  ISB1-2H14_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-2H14,
            genomic survey sequence.
ACCESSION   CL083840
VERSION     CL083840.1  GI:40558745
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
ORGANISM    Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1599)
AUTHORS     Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
TITLE       A physical map of the xenopus tropicalis genome
JOURNAL     Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000   Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence start: 1069
            High quality sequence stop: 1158.
FEATURES             Location/Qualifiers
     source          1..1599
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8364"
                     /clone="ISB1-2H14"
```

RESULT 12
CL038406/c

```
LOCUS       CL038406    1594 bp    DNA    linear   GSS 31-DEC-2003
DEFINITION  CH216-46A6_Sp6.1 CH216 Xenopus tropicalis genomic clone CH216-46A6,
            genomic survey sequence.
ACCESSION   CL038406
VERSION     CL038406.1  GI:40494319
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
ORGANISM    Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1594)
AUTHORS     Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
```

```
TITLE     Mardis,E. and Wilson,R.
          A physical map of the xenopus tropicalis genome
JOURNAL   Unpublished (2003)
COMMENT   Contact: Richard K Wilson
          Genome Sequencing Center
          Washington University School of Medicine
          Email: submissions@watson.wustl.edu
          Insert Length: 175000    Std Error: 0.00
          Seq primer: Sp6 ATTTAGGTGACACTATAG
          Class: BAC ends
          High quality sequence start: 471
          High quality sequence stop: 522.
FEATURES  Location/Qualifiers
  source  1..1594
          /organism="Xenopus tropicalis"
          /mol_type="genomic DNA"
          /strain="Nigerian frog"
          /db_xref="taxon:8364"
          /clone="CH216-46A6"
          /sex="male"
          /cell_line="Stock 248 F7A2, inbred N7"
          /clone_lib="CH216"
          /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
          BAC library"
ORIGIN

Query Match          6.0%;   Score 99;   DB 9;   Length 1594;
Best Local Similarity 45.8%;   Pred. No. 2.4e-08;
Matches 327;   Conservative 0;   Mismatches 387;   Indels 0;   Gaps 0;

Qy   384 TTGTATTTCCCTAGGATTAGGATTTGTTAGGGCAACGTTACGATTGCTCTGACAATA 443
         |||||||||   ||||||    ||||||      ||        ||
Db  1557 TTTTTTTTTTTTTTTATTATTTTTAATTTTTTTTTNTNTTTTTTTTTTTTTTT 1498

Qy   444 AATTAGAATTATTATTTTGTTACTTATGAGGTTATATCAACTTATGCCGACAATTTGTC 503
         |||||||    |||||||                    |||||||
Db  1497 TTTTTTTTTTTATTTTTTTTTTATTTTTTTTTTTTTTNTATT 1438

Qy   504 ATCGTAGTATTATATTCTGCAATTCCTCTTCTTCATTAGAAGTTATTATAGAAAATTATTT 563
         ||     |||||||        |||                     ||||
Db  1437 TTTTTATTTATTTTTTATTTTTTTTTTTTATATTATTTTNTTTTAT 1378

Qy   564 AATATTTCTAGTATTGAGAGATACACTGAAAGTTATTGTCCGTTTGTTGTTGT 623
         ||||            |||||||       ||||     ||||||||
Db  1377 TATTTTTTTTTTTTTATTTTTTTTTTTTTTTTATNTTTTTATT 1318

Qy   624 TTGTTTGTTTTTCAAAATATAGAATTACAAGAATATTAGTTGGGCGCTTTATTTGCAATA 683
         |||||||||    ||||||               ||||||||||
Db  1317 TTTTTTNATTTTTTTTTTTTTTTTTAATTTTTTTTTTTTTT 1258

Qy   684 AGTAATGTTGTAATAATGTACATTATGCAGTATATGCAGTATACCAATCTTGGATTGGACCTGTGTAAT 743
         |||||       |||||||           |||||      ||||||||||
Db  1257 TTTTTATTTTTTTTTTTTTATTTTTTTTTTTTATTATT 1198

Qy   744 TACTCCACTTGCATTTAAAGAAATTAATGAGATAATGCTGGCTAACAAATGATAGAT 803
         ||     |||||||              |||||       |||||
Db  1197 TTTTTTTATTTTTTTTTTTTTATTTTTTTTTTTTTTTTA 1138

Qy   804 AAATTCATATATCCATTGTTATTGGTTATTGGTTATTGAAGTTGCTGTGTGTTTTTGAAGTTTAAGT 863
         |||||        |||||||   |||||||          ||||||
Db  1137 TTTATTTTTTTTATTTTTTTATTTTTTTTTTTTTT 1078

Qy   864 TTCATAAAAGAAAAGTAATATAAAACTTCTGGATTTTGACTTTATTTTTATGCTGTG 923
         ||||||||          ||||||         |||||||||||
Db  1077 TTTTTTTTTTTATTTTTTTTATTTTTTTTTTTTTTTTT 1018

Qy   924 ATGATGTATGTTTTGTTCGAGCGTATACACAACAAAATCCCATGAGCGTTTTATTTCACCT 983
         ||      |||||||||            |||||||       |||||||||
Db  1017 TTTTTTTTATTTTTTTTTTTTATATTTTTTATATTTTTTTTTT 958

Qy   984 AACACTGTTTATTCCTCGATTAAAATCCAATTATTATCGTTGGGTTATTTTATAGGACGA 1043
         ||     ||||||        ||||||         ||||||      |||||||||
Db   957 TATTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTATTTTTTATAAATTTTT 898
```

```
Qy  1044 ATTGTTCCTTATGAGATATTTCTTTATCTAATATTCCTCTTATCATAAATCT 1097
         ||||   |||    ||||||||    ||||||||
Db   897 TTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTATTTTTTTTTATTT 844

RESULT 13
CL082000
LOCUS       CL082000  CH216-165P18_RM4.1 CH216 Xenopus tropicalis genomic clone
                      1981 bp    DNA    linear    GSS 31-DEC-2003
DEFINITION  CH216-165P18_RM4.1 CH216 Xenopus tropicalis genomic clone
            CH216-165P18, genomic survey sequence.
ACCESSION   CL082000
VERSION     CL082000.1   GI:40537913
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1981)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
  COMMENT   Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000    Std Error: 0.00
            Seq primer: RM4 ctcaagggcatcggtcgagc
            Class: BAC ends
            High quality sequence start: 265
            High quality sequence stop: 497.
FEATURES    Location/Qualifiers
  source    1..1981
            /organism="Xenopus tropicalis"
            /mol_type="genomic DNA"
            /strain="Nigerian frog"
            /db_xref="taxon:8364"
            /clone="CH216-165P18"
            /sex="male"
            /cell_line="Stock 248 F7A2, inbred N7"
            /clone_lib="CH216"
            /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis genome
            BAC library"
ORIGIN

Query Match          6.0%;   Score 98.6;   DB 9;   Length 1981;
Best Local Similarity 46.2%;   Pred. No. 2.8e-08;
Matches 326;   Conservative 0;   Mismatches 379;   Indels 0;   Gaps 0;

Qy   384 TTGTATTTCCTAGGATTTAGGATTTGTTAGGGCAACGTTACGATTGCTCTGACAATA 443
         |||||||||   ||||||   |||||||       ||        ||
Db   351 TTTTTTTATTTTTTTTTTTTTTTTTTTAATTTTTTTATTTTTTTTTTTTT 410

Qy   444 AATTAGAATTATTATTTTGTTACTTATGAGGTTATATCAACTTATGCCGACCAATTTGTC 503
         |||||||    |||||||                    |||||||
Db   411 TTTTTTTTTTATTTTTTTTTTTATTTTTTTTTTTTTTTTATTTTTTTTTTTT 470

Qy   504 ATCGTAGTATTATATTCTGCAATTCCTCTTCTTCATTAGAAGTTATTATTATAGAAAATTATTT 563
         ||     |||||||        |||                     ||||
Db   471 TTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTAAATTTTTTT 530

Qy   564 AATATTTCTAGTATTGAGAGATACACTGAAAGTTATTGTCCGTTTGTTGTTGT 623
         ||||            |||||||       ||||     ||||||||
Db   531 TTTTTTTTTTTATTTTTTTTTTATATTTTTTTTTTTTTTTTTTTTTTTT 590

Qy   624 TTGTTTGTTTTTCAAAATTATTTTTATTTTTTCAAAATTACAAGAATATTAGTTGGGCGCTTTATTTGCAAATA 683
         |||||||||    |||||||             ||||||||||
Db   591 TTTTTTTTTTTATTTTTTTTTTATTGTTTTTTTTTTTTTTTTTTTTTT 650

Qy   684 AGTAATGTTGTTAATAATGTACATTATGCAGTATATGCAGTACCAATCTTGGATTGGACCTGTTAAT 743
         |||||       |||||||           |||||      ||||||||||
```

```
Db   547 TNNNNNNNNNTTTTTTTTTNTTTTTTTTTTTTTNNTTTT 606
                                             
Qy   439 CAATAAATTAGAATTATTATTTTGTACTTATGAGGTATATCAACTTATGCGACAAT 498
             ||||| |||                    |||||
Db   607 NTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNNTTTTTTTTTTTTTTTTTNTTTT 666
                                        
Qy   499 TTGTCATCGTAGTATTATTCTGCAATTCTCTTCATTAGAAGTTATTATAGAAAAT 558
         ||| ||                          |||||||
Db   667 TTTTTTTTTTAATTTTTTTTTTTTTTGTTTTTTTTTTTTTNNTTTNTTNNTTTTT 726
                                        
Qy   559 TATTTAATATTCTAGTATTGAGAGATACACTGAAAGTTATTGTCGTTGTTGTTTG 618
             |||                         ||||||||
Db   727 TTTTTTTTTTTTTTTTTTTTTTTTATTCTTTTTTTTTTTTAFTTGTTATTTATT 786
                                        
Qy   619 TTTGTTGTTGTTTTTCAAAATATAGAATTACAAGAATTAGTGGGCCTTATTTG 678
         ||| ||                          ||||
Db   787 TTTTTTTTTTTTTTTTTTTTTTTTTNNNNTTTTNNTTTTTTTTTTTTTTTTGT 846
                                        
Qy   679 CAATAAGTATTGTGTTAATAATGTACATTATGCAGTATACCAATCTTGGATTGGACCTG 738
             ||                          ||||||
Db   847 TNTTTNNTTTNTTTNTTNTTTTTTTTTTTTNTNTTTTTTTTTTTTTTTTGTATT 906
                                        
Qy   739 TTAATTACTCACTGCATTTAAAGAAATTAATGAGATAACAAATCTGGCTAACAATGA 798
             ||                          |||||
Db   907 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNTNTTTTTTTTTTTTTN 966
                                        
Qy   799 TAGATAAATTCATATATCCATTGTGTTATTTCGTTTATTGAAGTTGCTGTGTTTTTAAGTT 858
             ||                          ||||||
Db   967 TTNNTNTTTTTTTTTTTTTTTTTTTTTTTNNTTTNTTTTNNTTTTTTTTTTTTTTTT 1026
                                        
Qy   859 TAAGTTTCATAAAAGAAAAGTATATAAACTTCCTGGGATTTTTGACTTTATTTTTATG 918
             ||                          ||||||
Db   1027 TTTTTTTTTNNNTTTTTNNNTTTTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1086
                                        
Qy   919 CTGTGATGATGTATGTTTTTGTTCGAGCGTATACACAAAATCCCATGAGCGTTTATTT 978
             ||                          ||||||
Db   1087 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNTNTTTTTTTTTTTNTTTTNNNTTTTTT 1146
                                        
Qy   979 CACCTAACACTGTTTATTCTCGATTAAAATCCAATTATTATCGTTGGGTATTTATAG 1038
             ||                          ||||||
Db   1147 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1206
                                        
Qy   1039 GACGAATTGTTCCTATGAGATATTTCTTATCCAATAATTCCTCCTTATCATAAATCT 1097
             ||                          ||||||
Db   1207 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNTTTT 1265
```

```
Db   651 TTTTTTTTTTTAGTTTTTTATTATTTTTTTATTTTTTTTATTATTTTTTTTTTT 710
                                        
Qy   744 TACTCACTTGCATTTAAAGAAATTAATGAGATAACAAATGCTGGCTAACAATGATAGAT 803
             ||||| |||                    |||||
Db   711 TTTTATTATTTTTTTTTTGTTTTTTTTATTTTTTTTTATTTTTATTT 770
                                        
Qy   804 AAATTCATATACCATTGTTATTGGTTTATTTGAAGTTGCTGTGTTTTTTAAGT 863
             ||||| |||                    |||||
Db   771 TTTTTTTTAATTTTTTTTATTGTTTTTGTTTTTTTTTTTTTATTTTTTTTTTT 830
                                        
Qy   864 TTCATAAAAGAAAAGTATATAAACTTCTCTGGATTTTTGACTTATTTTTATGTGCTGTG 923
             ||||| |||                    |||||
Db   831 TTTTTTTTTTTTTTATTTTTTTTATTCTTTTTTTTTTTTAFTTGTTATTTATT 890
                                        
Qy   924 ATGATGTATGTTTTGTTCGAGCCGTATACAACAAAATCCCATGAGCCGTTTATTTCACCT 983
             ||||| |||                    |||||
Db   891 TTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGT 950
                                        
Qy   984 AACACTGTTTATTCTCGATTAAAATCCAATTATTATCGTTGGGTATTTATAGGACGA 1043
             ||||| |||                    |||||
Db   951 ATTTTTTTTTTATTTATTATATATTTTTTTTTATTTTTTTTATTGTATT 1010
                                        
Qy   1044 ATTGTTCCTATGAGATATTTTCTTATCCAATAATTCCTCTTTAT 1088
             ||||| |||                    ||||||
Db   1011 TTTTTTTTTATTTTTTATTTTTTTTTTTTTTTTTTTTT 1055
```

RESULT 14
CG748432

LOCUS   CG748432   1344 bp   DNA   linear   GSS 24-OCT-2003
DEFINITION P042-2-C11.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG748432
VERSION   CG748432.1 GI:37969358
KEYWORDS   GSS.
SOURCE   Pristionchus pacificus
  ORGANISM Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 1344)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE   An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
  PUBMED   12884007
COMMENT   Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 0049707161371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES   Location/Qualifiers
     source   1..1344
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

Query Match   5.9%; Score 98.4; DB 9; Length 1344;
Best Local Similarity 41.3%; Pred. No. 3e-08;
Matches 297; Conservative 0; Mismatches 422; Indels 0; Gaps 0;

```
Qy  379 TAGAATTGTATTTCCTAGGATTTAGGATTTGTTGGTAGGGCAACGTTACGATTGCTCTGA 438
            ||  |||||     || ||||||         ||||| ||
```

RESULT 15
CL082569/c

LOCUS   CL082569   1632 bp   DNA   linear   GSS 31-DEC-2003
DEFINITION CH216-167P5_Sp6.1 CH216 Xenopus tropicalis genomic clone
            CH216-167P5, genomic survey sequence.
ACCESSION   CL082569
VERSION   CL082569.1 GI:40538482
KEYWORDS   GSS.
SOURCE   Xenopus tropicalis (western clawed frog)
  ORGANISM Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1 (bases 1 to 1632)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE   A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 1062
```

Job time : 5402 secs

```
FEATURES    High quality sequence stop: 1114.
source      Location/Qualifiers
            1..1632
                   /organism="Xenopus tropicalis"
                   /mol_type="genomic DNA"
                   /strain="Nigerian frog"
                   /db_xref="taxon:8364"
                   /clone="CH216-167P5"
                   /sex="male"
                   /cell_line="Stock 248 F7A2, inbred N7"
                   /clone_lib="CH216"
                   /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                   BAC library"

ORIGIN

Query Match          5.9%;   Score 98.4; DB 9; Length 1632;
Best Local Similarity 45.7%; Pred. No. 3e-08;
Matches 353;  Conservative  0; Mismatches 418; Indels   1; Gaps   1;

Qy      328 AATAGTTAGAAATCTTGTTGAATTTAATTAGATAAAAAATATTGTACAGGGTAGAATTGT 387
                |||||||||||||||||  |  | | |||||  ||||||
Db     1608 ATTATTTTTTTATATTTTATTTTTTTTTTTTTTTTTTTTTTTATTNATTTTTTTTTT 1549

Qy      388 ATTTTCCTAGGATTAGGATTTGTTGAGGGCAACGTTACGATTGCTCTGACAATAAATT 447
                |||||||      |||||||       ||||||
Db     1548 TTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTATNTTTTNTTTTTTT 1489

Qy      448 AGAATTATTATTTTTGTTACTTTATGAGGTTATATCAACTTATGCGACAATTTGTCATCG 507
                ||  ||  ||||||||||||           ||||||
Db     1488 TATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTATTTT 1429

Qy      508 TAGTATTATATTCTGCAATTCTTCTTCATTAGAAGTTATTTATAGAAAATTATTTAATA 567
                ||||||||||||||||||        |||||||
Db     1428 TATTTTTTTTTTTTTTNATTNTTTTTATTNTTTTTTTTTTTTTTTTTATTT 1369

Qy      568 TTTCTAGTAGTGAGAGATACACTGAAAGTTATTCGTCTGCTTTGTT-TGTTTGTTGTTTG 626
                |||||||||||||||||||||              | | |||
Db     1368 TTTTTTTTTTTTTTTTTTTTTTTATNTNATTTTTTNTTTTTTTTTTT 1309

Qy      627 TTTGTTTTTCAAAATATAGAATTACAAGAATATTAGTTGGCGCTTATTGCAATAAGT 686
                ||||||      |||||       |||||||
Db     1308 TATTTTTTTTTTTATTTTTTTTNATTTTTTATTTTTTTTAATTTTTTTTTTT 1249

Qy      687 ATTGTTGTTAATAATGTACATTATGCAGTATACCAATCTTGGATTGGACCTGTTAATTAC 746
                ||||||           |||||||||||||||
Db     1248 TTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTATATTTTTTTATTTN 1189

Qy      747 TCACTTGCATTAAAGAAAAATTAAATGAGAATAACAAATGCTGGCTTAACAATGATAGATAAA 806
                |||||||||||                 ||||||
Db     1188 TNTTATTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTAATTATTTATTTT 1129

Qy      807 TTCATATCCATTGTTATTGGTTATTCGAAGTTGCTGTGTTTTAAGTTAAGTTTC 866
                |||||||||||||||       ||||||   |||
Db     1128 TTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1069

Qy      867 ATAAAAGAAAAGTAATAAAACTTTCTTGGATTTTGACTTATTTTTATGCTGTGATG 926
                |||||||||||||||||      ||||||   |||
Db     1068 TTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1009

Qy      927 ATGTATGTTTTGTTCGAGGCTATACAACAAACATCCCATGAGCGTTTATTTCACCTAAC 986
                ||||||||||||||          ||||||   |||
Db     1008 TTTTTTTTTTTTTTTATTTTTTTTTTTTAATATATTTTTTAATATTTTTTT 949

Qy      987 ACTGTTTATTCTCGATTAAAATCCAATTATTTATCGTTGGGTTATTTATAGGACGAATT 1046
                |||||||||||||||       |||||||   |||
Db      948 TTTTTTTTTTTTTTATTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTT 889

Qy     1047 GTTCCTTATGAGAATATTTTCTTATCTAATATTCCTCTTTATCATAAATCTA 1098
                ||  ||||||||||||                ||||||
Db      888 ATTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTATA 837
```

Search completed: December 23, 2004, 19:46:21